Jiang, S.Y., Zhang, Q., Werner, A.D., Wellen, C., **Jomaa, S.**, Zhu, Q.D., **Büttner, O.**, Meon, G., **Rode, M.** (2019):
Effects of stream nitrate data frequency on watershed model performance and prediction uncertainty
J. Hydrol. **569** , 22 – 36

http://dx.doi.org/10.1016/j.jhydrol.2018.11.049

Elsevier Editorial System(tm) for Journal of Hydrology
Manuscript Draft

Manuscript Number: HYDROL29221R1

Title: Effects of stream nitrate data frequency on watershed model performance and prediction uncertainty

Article Type: Research paper

Keywords: Nitrate export; HYPE; Monitoring frequency; Model calibration; Prediction uncertainty; DREAM

Corresponding Author: Dr. Sanyuan Jiang,

Corresponding Author's Institution: Nanjing Institute of Geography and Limnology,Chinese Academy of Sciences

First Author: Sanyuan Jiang

Order of Authors: Sanyuan Jiang; Qi Zhang; Adrian Werner; Christopher Wellen; Seifeddine Jomaa; Qiande Zhu; Olaf Büttner; Meon Günter; Michael Rode

Abstract: High-frequency water quality monitoring is increasingly used in examining the nutrient fluxes within catchments. Despite this, no studies have assessed the impact of monitoring frequency on the uncertainty of nitrate estimates obtained from distributed or semi-distributed catchment models. This study aims to evaluate the impacts of two different frequencies of nitrate sampling on the performance of a catchment hydrology model, including the uncertainty in both predictions and calibrated parameters. The investigation uses the HYPE model to simulate streamflow and nitrate concentrations (2010-2015) in the Selke catchment, a heterogeneous mesoscale catchment in central Germany. The Bayesian inference scheme of the DREAM code was employed for calibration and uncertainty analysis, and to explore differences between fortnightly and daily nitrate sampling strategies. The results indicate that: (a) the posterior uncertainty intervals of nitrogen-export process parameters were narrower when the model was calibrated to daily nitrate measurements, while similar maximum likelihood parameter values were obtained regardless of the sampling frequency; (b) the model calibrated using daily nitrate data better represented both daily and fortnightly nitrate measurements relative to that obtained using fortnightly sampling; (c) the daily nitrate dataset produced significantly smaller parametric prediction uncertainty, but only modest reduction in total prediction uncertainty, relative to the fortnightly nitrate dataset; (d) model structural error and measurement errors are the primary sources of total prediction uncertainty, and these combine to inhibit the benefits of high-frequency monitoring. We conclude that the adequacy of sampling frequency is dependent on model structure and measurement errors, such that higher-frequency nitrate monitoring may not markedly reduce the uncertainty of nutrient predictions, depending on other levels and sources of uncertainty.

Highlights (3 to 5 bullet points (maximum 85 characters including spaces per bullet point)

Highlights:

- Nitrate export uncertainty in Selke catchment model examined using HYPE and DREAM
- Posterior parameter uncertainty decreased using daily than fortnightly nitrate data
- Daily-calibrated model better captured nitrate dynamics than using fortnightly data
- Daily nitrate data produced smaller predictive uncertainty than fortnightly data
- Model structural- and input errors inhibit the benefits of high-resolution data

Effects of stream nitrate data frequency on watershed model performance and prediction uncertainty

S. Y. Jiang[1,2,3*], Q. Zhang[1], A. D. Werner[4], C. Wellen[5], S. Jomaa[3], Q. D. Zhu[2], O. Büttner[3], G. Meon[6], M. Rode[3]

[1]Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China

[2]State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing 210029, China

[3]Department of Aquatic Ecosystem Analysis and Management, Helmholtz Centre for Environmental Research-UFZ, Magdeburg 39114, Germany

[4]College of Science and Engineering, and National Centre for Groundwater Research and Training, Flinders University, GPO Box 2100, SA 5001, Australia

[5]Department of Geography and Environmental Studies, Ryerson University, Toronto, 350 Victoria Street Ontario, Canada

[6]Department of Hydrology, Water Management and Water Protection, Leichtweiss Institute for Hydraulics and Water Resources, University of Braunschweig, Braunschweig 38106, Germany.

[*]Corresponding author: Sanyuan Jiang (syjiang@niglas.ac.cn)

Originally submitted to *Journal of Hydrology* on 7 April 2018
Revision submitted to *Journal of Hydrology* on 6 September 2018
Revision submitted to *Journal of Hydrology* on 2 November 2018

Abstract

High-frequency water quality monitoring is increasingly used in examining the nutrient fluxes within catchments. Despite this, no studies have assessed the impact of monitoring frequency on the uncertainty of nitrate estimates obtained from distributed or semi-distributed catchment models. This study aims to evaluate the impacts of two different frequencies of nitrate sampling on the performance of a catchment hydrology model, including the uncertainty in both predictions and calibrated parameters. The investigation uses the HYPE model to simulate streamflow and nitrate concentrations (2010-2015) in the Selke catchment, a heterogeneous mesoscale catchment in central Germany. The Bayesian inference scheme of the DREAM code was employed for calibration and uncertainty analysis, and to explore differences between fortnightly and daily nitrate sampling strategies. The results indicate that: (a) the posterior uncertainty intervals of nitrogen-export process parameters were narrower when the model was calibrated to daily nitrate measurements, while similar maximum likelihood parameter values were obtained regardless of the sampling frequency; (b) the model calibrated using daily nitrate data better represented both daily and fortnightly nitrate measurements relative to that obtained using fortnightly sampling; (c) the daily nitrate dataset produced significantly smaller parametric prediction uncertainty, but only modest reduction in total prediction uncertainty, relative to the fortnightly nitrate dataset; (d) model structural error and measurement errors are the primary sources of total prediction uncertainty, and these combine to inhibit the benefits of high-frequency monitoring. We conclude that the adequacy of sampling frequency is dependent on model structure and measurement errors, such that higher-frequency nitrate monitoring may not markedly reduce the uncertainty of nutrient predictions, depending on other levels and sources of uncertainty.

Keywords: Nitrate export; HYPE; Monitoring frequency; Model calibration; Prediction uncertainty; DREAM

1. Introduction

Excessive nutrient export from landscapes to surface water bodies (e.g., rivers, reservoirs and lakes) has caused water quality deterioration and eutrophication in aquatic environments across the globe (Conley et al., 2009). The management of nutrient sources and transport pathways within catchments is increasingly based on catchment hydrological and biogeochemical modelling (Rode et al., 2010; Shrestha et al., 2012; Wellen et al., 2015). Commonly used hydrological and nutrient transport models include SWAT (Soil and Water Assessment Tool; Arnold et al., 1998), AGNPS/AnnAGNPS (Agricultural Non-Point Source Pollution Model; Bingner et al., 2012), INCA (Integrated Nutrients from Catchment; Whitehead et al., 1998), HSPF (Hydrological Simulation Program-Fortran; Bicknell et al., 2012) and HYPE (HYdrological Predictions for the Environment; Lindström et al., 2010). Most off-the-shelf watershed-scale hydrologic and nutrient models currently used for decision-making and management operate at a daily time step. Catchment hydrological models rely on a mixture of empirically and physically based parameters. Empirical parameters are difficult to compare to measurable field parameters, but are nonetheless required to represent complex flow and transport processes that are otherwise challenging to simulate using measurable parameters, in particular at the catchment scale. These are routinely assigned values based on model calibration against field measurements (e.g., Liu and Gupta, 2007).

In heterogeneous catchments, solute fluxes (e.g., dissolved inorganic nitrogen) are affected by hydrogeological and landscape characteristics, such as flow pathways, transit times, lithology, land cover and soil wetness (Hrachowitz et al., 2016; Jiang et al., 2014; Onderka et al., 2012; van

Griensven et al., 2006; Viswanathan et al., 2016). The representation of these processes in catchment models depends on several factors, in particular, the spatial and temporal discretization. The spatial resolution of catchment hydrological and nutrient export models is often limited by the distribution of monitoring, in particular, the gauging of flow and nutrient concentrations in streams. Jiang et al. (2015) compared nutrient runoff models based on single-site or multi-site stream gauging stations, for the Selke River (referred to locally as simply "Selke") catchment (Germany). They found that multi-site calibration improved the performance of their nitrate model, indicated by an increase of 12% in the Nash-Sutcliffe coefficient (for calibration mismatch), and a decrease of 33% in the uncertainty of nitrate predictions. Similarly, Cao et al. (2006) found that multi-variable and multi-site calibration improved the performance of a hydrological model (SWAT) of the Motueka catchment (New Zealand), and helped to identify the areas and hydrological processes requiring greater scrutiny and calibration effort.

Nitrate export is often highly variable over time due to various factors, including the temporal variability in the hydrological regime, nutrient inputs, and biogeochemical processes (Basu et al., 2011; Li et al., 2010; Rode et al., 2016a; Molenat et al., 2008; van Griensven et al., 2006; Van Meter et al., 2016; Van Meter and Basu, 2015; Viswanathan et al., 2016). The high temporal variability in water quality has led to the development of high-frequency (e.g., up to daily intervals) measurement capabilities. Establishing and assessing sampling frequencies is necessary in order to yield information needed for policy making and catchment management plan implementation, and to ensure the cost effectiveness of water quality monitoring plans (Behmel et al., 2016). Stream water quality measurements at different temporal resolutions (e.g., daily, weekly, fortnightly, monthly) have been utilized in previous analyses of water quality variations, nutrient loads, nutrient sources, transport pathways and retention processes within catchments (e.g., Ullrich and Volk, 2010; Wade et al., 2012; Yang et al., 2018). Numerous water quality monitoring studies, from a variety of different catchments, illustrate that lower-frequency (e.g., monthly) nitrate monitoring may characterize reasonably well the average nitrate conditions, but often fails to capture short-term water quality dynamics and extremes (Fovet et al., 2015; Halliday et al., 2015; Jones and Chappell, 2014; Ross et al., 2015; Wade et al., 2012). This is only apparent when higher-frequency sampling is available. Based on statistical uncertainty analysis and using long-term nitrate measurements, Levine et al. (2014) demonstrated that uncertainty in the detection of changes in stream nitrate concentrations over time increased when the sampling frequency was reduced from weekly to bimonthly.

Previous catchment modelling investigations have adopted observations of stream nutrient concentrations at daily to monthly intervals in calibrating and validating nutrient export models (e.g., Ahmad et al., 2011; Jomaa et al., 2016; Lam et al., 2012; Lindström et al., 2010; Pathak et al., 2018; Shrestha et al., 2007; Woodward et al., 2017). A limited number of nitrate export modelling studies have assessed different frequencies of nitrate monitoring adopted in calibration. For example, Woodward et al. (2017) showed that the simple lumped-parameter, daily time-step catchment model StreamGEM predicted similar nitrate dynamics when calibrated using either monthly nitrate data or daily nitrate data. Several studies have similarly concluded that nitrate export modelling based on long-term, monthly nitrate datasets potentially represent reasonably well the seasonal dynamics of stream nitrate concentrations and loads (Jiang et al., 2014; Pathak et al., 2018; Woodward et al., 2017). The general conclusion arising from these investigations is that monthly sampling may be sufficient to support water resource management decision-making.

However, other studies demonstrate that low-frequency water quality sampling may restrict the calibration (and validation) of catchment nutrient export models (Chappell et al., 2017; Kirchner et al., 2004; Pathak et al., 2018), although a proper investigation of model uncertainty has not been undertaken to support these claims. Previous evaluations of nutrient monitoring frequency are generally based on highly simplified models of catchment processes. More generally, daily nitrate data are rarely incorporated into semi-distributed process-based watershed modelling studies (Jomaa et al., 2016; Lam et al., 2012). Ahmad et al. (2011) illustrated that calibration of their SWAT model using measurements at monthly intervals resulted in considerable underestimation of monthly sediment and nitrogen loads, most notably during intensive rainfall periods. Under-prediction of nitrate loads during storm flow conditions was also noted from a watershed nitrate transport model calibrated using daily nitrate data, due to the underestimation of peak streamflow rates (Lam et al., 2012). It appears from these studies that higher-frequency monitoring has greater opportunity to avoid sampling-related artefacts that arise when measurements are used in parameterizing process-based watershed models used for water quality simulation (Jones et al., 2014).

To date, the use of high-frequency (e.g., up to daily) nutrient measurements in watershed water quality modelling is limited (Jones et al., 2014). High-frequency hydrochemical monitoring has mainly been used for: (a) the direct observation and assessment of water quality variations (Halliday et al., 2015), (b) the direct estimation of pollutant loads, calculated as the product of measured flow and solute concentrations (Ullrich and Volk, 2010; Wade et al., 2012; Jomaa et al., 2018), (c) the analysis of river pollutant sources and transport pathways (Aubert et al., 2013; Halliday et al., 2015; Lloyd et al., 2016; Wade et al., 2012), and (d) the analysis of in-stream biogeochemical processes (e.g., Rode et al., 2016a, Wade et al., 2012). High-frequency in-situ monitoring allows short-term water quality dynamics to be captured, leading to important knowledge gains of catchment hydrochemical sources and behavior (e.g., in-stream primary production) that may otherwise be unattainable from regular low-frequency (e.g., fortnightly, monthly) monitoring (Halliday et al., 2015; Rode et al., 2016b; Sandford et al., 2013; Skarbøvik et al., 2012). The benefits of adopting high-frequency sampling strategies have been demonstrated for the direct calculation of nutrient and sediment loads. For example, Skarbøvik et al. (2012) reported that the reliability of average concentrations of suspended particulate matter improved with decreasing sampling interval (i.e., from monthly to daily), such that sampling at monthly intervals resulted in underestimation of sediment loads of up to 98%. This arises because fortnightly-to-monthly sampling frequencies may miss significant nutrient and/or sediment fluxes during storm events, resulting in large errors in load estimation, especially given that high-flow conditions typically produce the highest nutrient and sediment transfer rates (Horowitz, 2003; Jordan and Cassidy, 2011; Rodríguez-Blanco et al., 2013; Sharpley et al., 2008).

Where nutrient and sediment fluxes are determined from catchment-scale hydrological models, uncertainty arises in the model predictions due to several factors (Ajami et al., 2007; Balin et al., 2010; Beven and Binley, 1992). These include measurement errors in model input data (e.g., rainfall, diffuse and point nutrient sources, land use, etc.) and calibration data (e.g., flow rates, nutrient concentrations, etc.), uncertainty in model parameters that persists after model calibration (i.e., posterior parameter uncertainty), and model structural errors. Higher posterior parameter uncertainty is associated with parameters that are non-uniquely estimated by model calibration, and these tend to be weakly identifiable from the calibration dataset (e.g., Knowling and Werner, 2016). Model structural errors include the averaging associated with the spatial and temporal resolution of calculations (i.e., relative to field-scale variability), and the intrinsic inability of a given model to reproduce the physical and biogeochemical mechanisms involved in runoff generation and nutrient export (e.g., Ajami et al., 2007; Balin et al., 2010; Rode et al., 2010; Vrugt, 2016; Woodward et al., 2017; Yang et al., 2007). We refer herein to the summation of model structural errors and measurement errors as "residual errors". The uncertainty that arises from these unavoidable limitations of hydrological models manifests as posterior parameter uncertainty, and uncertainty in nutrient and flow predictions (i.e., "prediction uncertainty"). Prediction uncertainty (referred to hereafter as "total prediction uncertainty") has two components: residual error and parametric prediction uncertainty; the latter being the uncertainty in model simulation results that arises out of posterior parameter uncertainty. The current study assesses both total prediction uncertainty and parametric prediction uncertainty, thereby allowing us to infer residual error as the difference between the two. Posterior parameter uncertainty is also evaluated.

Uncertainty analysis is a requisite component of catchment hydrology modelling activities because model parameters are often poorly constrained by field evidence and rely heavily on calibration, and because parameter interdependence may lead to non-unique predictions of future catchment behavior (Beven and Binley, 1992; Wellen et al., 2015). Uncertainty analysis also allows for evaluation of the information content of different observations, thereby guiding field measurement campaigns and strategies for model development (Arhonditsis et al., 2007; Rode et al., 2010). Additionally, the provision of prediction uncertainty intervals allows for risk-based decision making within watershed management (Arhonditsis et al., 2007). Uncertainty analysis is particularly important in the application of catchment hydrology models to estimate nutrient fluxes (Kyllmar et al., 2014). Some popular uncertainty analysis techniques used in catchment runoff and nutrient export modelling include GLUE (Generalized Likelihood Uncertainty Estimation; Beven and Binley, 1992; Gong et al., 2011), SUFI-2 (Sequential Uncertainty Fitting algorithm; Abbaspour et al., 2004; Wu and Chen, 2015), SCEM-UA (Shuffled Complex Evolution Metropolis algorithm; Dotto et al., 2012; Vrugt et al., 2003) and DREAM (DiffeRential Evolution Adaptive Metropolis algorithm; Jiang et al., 2015; Laloy and Vrugt, 2012). The Bayesian inference algorithm of DREAM has been widely used and proven to be appropriate to assess model uncertainty for a range of modelling applications. It is able to constrain both total prediction uncertainty, parametric prediction uncertainty and posterior parameter uncertainty (Jiang et al., 2015; Woodward et al., 2017; Yang et al., 2008). Moreover, DREAM is efficient for high-dimensional searches (i.e., exploring posterior distributions of a large number of parameters in complex models) for global minima in model error, in attempting to optimize model parameters (Laloy and Vrugt, 2012; Vrugt, 2016).

While uncertainty analysis is commonly conducted on rainfall-runoff modelling to assess streamflow prediction uncertainty caused by various factors, it is rarely implemented in nutrient export modelling studies (Jiang et al., 2015; Pathak et al., 2018; Wellen et al., 2015; Woodward et al., 2017). Examples include the analysis of uncertainty in catchment phosphorous export estimates from SWAT modelling, using GLUE (e.g., Gong et al., 2011), the Bayesian assessment of uncertainty in total phosphorus flux predictions from SPARROW modelling (e.g., Kim et al., 2017), and the Bayesian evaluation of model structure in the performance and uncertainty of four

Export Coefficient Models used to simulate nitrate export (Xia et al., 2016). Woodward et al. (2017) illustrated that daily nitrate data, used in calibration, did not reduce the uncertainty of nitrate export predictions of the StreamGEM model compared to monthly nitrate data. However, they used a simple, lumped-process model that assumed constant nitrate concentrations discharged from near-surface runoff, fast groundwater and slow groundwater flow paths, for nitrate export simulation. To our knowledge, no previous nitrate export studies that adopt process-based, semi-distributed watershed models have attempted to assess the uncertainty of parameters and predictions related to various sampling frequencies (e.g., fortnightly, daily) in nitrate measurements.

The objective of this study is to evaluate in a comprehensive manner the posterior parameter uncertainty, calibration mismatch, total prediction uncertainty, parametric prediction uncertainty, and residual error (i.e., the sum of measurement error and model structural error) associated with a watershed nitrate export model that arise from different frequencies (daily, fortnightly) of in-stream nitrate measurements. A case study of the Selke catchment is used as the basis for this evaluation. DREAM (version $DREAM_{ZS}$; Laloy and Vrugt, 2012) was used to calibrate the HYPE model and assess the uncertainty of parameters and model outputs. The current investigation assesses, for the first time, the effects of different frequencies (fortnightly or daily) of nitrate measurements in the simulation of nitrate export using a catchment-scale distributed-parameter model. In doing so, we expect that the findings will assist in designing other watershed models aimed at nutrient export estimation, and in developing water quality monitoring programs (i.e., establishing appropriate monitoring frequencies) that aim to inform nitrate export models and water resource management more generally.

2. Materials and Methodology

2.1 Study site and data

A heterogeneous nested mesoscale catchment (Selke catchment, central Germany) was selected as the study area for testing different sampling frequencies. The current study builds on previous investigations of the Selke catchment by a subset of the current author list. That is, Jiang et al. (2014) evaluated the capability of the HYPE model to represent spatial and temporal variability in nitrogen fluxes, while Jiang et al. (2015) assessed the impact of the spatial resolution in nitrogen observations on nitrogen export modeling, again using the HYPE model. The process-based HYPE model was utilized again in the current investigation to simulate streamflow and stream water nitrate concentrations. Among the available codes, the HYPE model requires readily available meteorological data (rainfall, temperature), and offers apposite balance between representation of hydrological, nutrient-export processes and model complexity. It has been applied successfully to the simulation of streamflow and water quality for catchments with a wide range of meteorological, hydrological and physiographic characteristics, but importantly, this includes catchments that have features similar to the study area adopted for the current investigation (e.g., Jiang et al., 2014; Lindström et al., 2010; Yin et al., 2016).

Selke is a tributary of the Bode River, which originates in the southwest of the state of Saxony-Anhalt (central Germany), in the vicinity of Harz Mountain. Selke drains an area of 463 km^2, with elevations varying between 53 m and 605 m (Figure 1). The Selke catchment is dominated by forest (broad-leaved, coniferous and mixed forests) in mountainous areas, and by agriculture in lowland areas; these accounting for 35% and 52% of the total catchment area, respectively. The remaining 13% of the catchment is used for pastures and urban land. Soil types are dominated by cambisols (schist and claystone) in the mountainous areas and chernozems (tertiary sediments with loess) in the lowland areas. Average precipitation decreases from 792 mm/yr in the mountainous areas to 450 mm/yr in the lowland areas, with an average of 660 mm/yr for the whole catchment. Annual precipitation is summer dominant, with a ratio between summer and winter of 1.35. The average temperature is 9°C, with an average monthly low of -1.8°C in January and monthly high of 15.5°C in July. Crops mainly consist of winter wheat, triticale, winter barley, rye, rapeseed and corn. Fertilizer inputs range from 130 to 190 kg N/ha/yr for nitrogen, and from 20 to 30 kg P/ha/yr for phosphorus (Kistner et al., 2013). Point source inputs from sewage plants contribute only a limited share of the total nitrogen load (about 2.9%), and more than 95% of the households are connected to public wastewater treatment plants (Rode et al., 2016a).

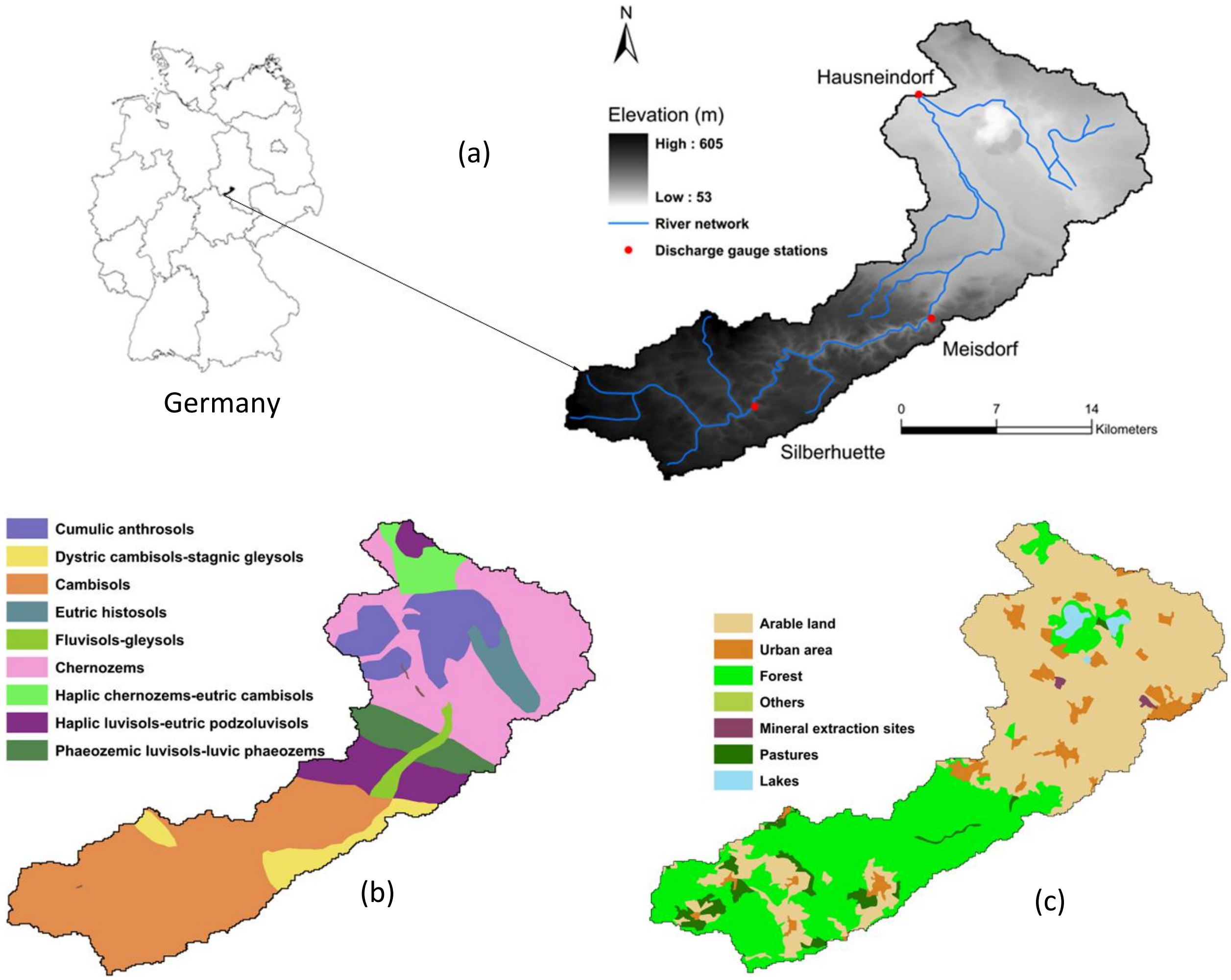

Figure 1. The Selke catchment: (a) Digital Elevation Model (DEM) and locations of streamflow and water quality gauging stations, (b) soil types, and (c) land use.

An extensive dataset is available to develop and calibrate a watershed model of the Selke catchment. A Digital Elevation Model (DEM) and soil-type distributions were obtained from the state survey office with grid resolutions of 90 m and 50 m, respectively. Land use was interpreted from Corine Land Cover 2006 of Germany (https://land.copernicus.eu/pan-european/corine-land-cover) with grid resolution of 25 m. The Selke catchment represents one of the best meteorologically and hydrologically equipped catchments in central Germany, with long-term monitoring of precipitation, temperature, discharge, water quality, etc. There are 16 rainfall stations and two climate stations within or close to the Selke catchment, although only four rainfall stations have been running since 2010. Daily rainfall data were provided by the German Weather Service. Stream discharge has been measured at three gauging stations, namely Silberhuette (upstream), Meisdorf (middle-stream) and Hausneindorf (catchment outlet), since 1920 (see Figure 1).

Within the TERENO (TERrestrial Environmental Observatories, http://teodoor.icg.kfa-juelich.de/overview-en?set_language=en) program, continuous high-frequency stream nitrate concentrations have been measured in-situ at stream gauging stations at 15-min intervals since 2010 using TRIOS ProPS-UV sensors with an optic path length of 10 mm (Rode et al., 2016a; Wollschläger et al., 2017). Maintenance of the instruments and calibration of sensors are conducted fortnightly. During instrument maintenance, stream water samples are taken for laboratory analysis to validate in-situ stream nitrate measurements, and to measure other hydrochemical constituents that are not recorded by field instruments. The measured nitrate concentrations obtained from laboratory analysis of fortnightly samples were used as the fortnightly nitrate dataset in this study. The in-situ nitrate concentrations were highly consistent with laboratory analysis data, evidenced by an R^2 (coefficient of determination) value of 0.93 (Rode et al., 2016a). The daily nitrate dataset adopted in the current model development consisted of arithmetic averages (for each day) of 15-min stream nitrate observations.

The average streamflow rates (during 1994-2004) through the two most upstream stations (i.e., both situated in the mountainous region) were 1.30 m^3/s (Silberhuette) and 1.52 m^3/s (Meisdorf), while the streamflow rate at the lowland catchment outlet (Hausneindorf) was 1.75 m^3/s during the same period. These flow rates amount to area-weighted runoff (i.e., the specific discharge) from the catchments upstream of the Silberhuette, Meisdorf and Hausneindorf stations of 415 mm/yr, 265 mm/yr and 133 mm/yr, respectively, indicating that catchment runoff is significantly higher in upstream mountainous areas (Jiang et al., 2014). The average streamflow rates during 2010-2015 were 0.82, 1.45 and 1.59 m^3/s from the gauging stations of Silberhuette, Meisdorf and Hausneindorf, respectively. The corresponding area-weighted runoff values were 265 mm/yr, 241 mm/yr and 128 mm/yr. Taking the difference between average flow rates for the three gauging stations produces area-weighted runoff values for the sub-basin between Silberhuette and Meisdorf of 74.2 mm/yr, and for the sub-basin area between Meisdorf and Hausneindorf stations of 44.1 mm/yr. The considerable reduction in area-weighted runoff in downstream areas relative to upstream areas is attributable primarily to the rainfall gradient, as mentioned above, and the steeper topography of upland regions (Jiang et al., 2014).

The average stream nitrate concentrations were 1.44, 1.75 and 3.91 mg/L at Silberhuette, Meisdorf and Hausneindorf stations, respectively, during the period 1994-2004 (based on fortnightly sampling). Using the 15-min nitrate measurements during 2010-2015, the average stream nitrate concentrations were 1.55, 1.63 and 2.89 mg/L at Silberhuette, Meisdorf and Hausneindorf stations, respectively. Higher nitrate concentrations in lowland areas are likely caused by fertilizer applications and accompanying leaching into surface and subsurface pathways, which flow to watercourses.

2.2 Hydrological and nitrogen-export model

HYPE was developed by the Swedish Meteorological and Hydrological Institute based on the hydrological-nutrient model HBV-NP (Lindström et al., 2010). It has been widely used to simulate streamflow and nutrient export, and to assess the impacts of agricultural practices on nutrient yields at catchments of different scales, and with various meteorological, hydrological and physiographical characteristics (Jiang et al., 2014; Jiang et al., 2015; Jomaa et al., 2016; Lindström et al., 2010; Strömqvist et al., 2012; Yin et al., 2016). Typical application of the HYPE model involves firstly delineating the study catchment into sub-basins based on the DEM and stream network. Each sub-basin is divided into Soil-Land-use Classes (SLCs) by overlaying maps of land use and soil type, with each SLC corresponding to a so-called Hydrological Response Unit (HRU). SLCs are considered as the smallest computational spatial units. The SLCs are not coupled to geographic locations, but defined as fractions of a sub-basin area. In each SLC, soil is divided vertically into one or several (maximum three) layers, which may have different thicknesses. Hydrological and nutrient processes are simulated within each soil layer of each SLC. A detailed description of the model structure, equations and parameters of HYPE are given by Lindström et al. (2010), and are therefore not repeated in detail here. Only a summary of the model structure, hydrological and nitrogen processes, parameter values, and the methodology of applying HYPE are given below.

HYPE simulates a wide range of hydrological processes, including snow accumulation and snowmelt, evapotranspiration, surface runoff, infiltration, macro-pore flow, percolation, soil runoff, tile-drain flow, regional groundwater flow, and river flow (Lindström et al., 2010). The water transit time is equal to the storage volume of a flow component (e.g., interflow, regional groundwater flow) divided by the outflow from that component (see eqs. (1) to (3) below). Lindström et al. (2010) and Tonderski et al. (2017) found that HYPE reproduced to a reasonable accuracy the temporal dynamics of groundwater levels and ^{18}O isotope concentrations in both forested and agricultural catchments, verifying the model's representation of water pathways and transit times. The latter were found to be in the order of months to years. The following equations link runoff rates to transit times in HYPE.

$$q_{RUNF} = \partial_{RC} \cdot (W_{SOIL} - \alpha_{FC}) \quad (1)$$

$$Q_{RUNF} = \sum_{\forall land\ class} q_{RUNF} \cdot c_{AREA} \quad (2)$$

$$T = \frac{\sum_{\forall land\ class}(W_{SOIL} - \partial_{FC}) \cdot c_{AREA}}{Q_{RUNF}} \quad (3)$$

Where q_{RUNF} is soil runoff (per area) from an SLC (mm/d), Q_{RUNF} is soil runoff from a sub-basin (mm m^2/d), T is water transit time (d), ∂_{RC} is the soil runoff coefficient (1/d), W_{SOIL} is soil moisture storage (mm), α_{FC} is the water content at the threshold for runoff (mm), and c_{AREA} is the SLC area (m^2).

Within HYPE, sources of nutrient input to the soil include diffuse sources from applied organic and inorganic fertilizer, manure, plant residues, atmospheric deposition, and rural households, and point sources from urban and industrial activities (e.g., sewage treatment works). The simulated stored mass of soil nitrogen consists of a fastN pool (i.e., stored mass of organic nitrogen in the soil with rapid transformation to dissolved organic nitrogen and inorganic nitrogen), a humusN pool (i.e., stored mass of organic nitrogen in the soil with a slow transformation to fastN), an organicN pool (i.e., stored mass of organic nitrogen in the soil available for mineralization to inorganicN) and an inorganicN pool (i.e., stored mass of inorganic nitrogen in the soil). The simulated nitrogen transformation processes are degradation, mineralization, denitrification and plant uptake within the soil profile, and denitrification, mineralization, and primary production in the river. Nitrogen transport follows the same pathways as water in the model. All nitrogen in the inorganic nitrogen pool is considered to be mobile and can hence be transported between soil layers or out of the profile through horizontal and lateral soil water flow, and regional groundwater flow. All biogeochemical processes in the soil, regional groundwater, streams and rivers are calculated by empirical equations that include first-order reaction rates, nutrient concentrations within each storage pool (e.g., humusN, fastN, organicN, inorganicN, total phosphorus) and influential environmental factors, such as soil water content, soil temperature, water temperature, and surface water volumes and surface areas (Lindström et al., 2010).

The hydrological and nitrogen-export processes, and the parameters for soil and surface water components in the HYPE model are listed in Table 1. Table 1 also includes parameter ranges considered plausible for the conditions encountered in Selke. These are based on knowledge of the relevant hydrological and nutrient processes, literature review, and previous HYPE applications to other German and Swedish catchments, such as Weida, Rönneå and Vindån (e.g., Jiang et al., 2014; Jiang et al., 2015; Jomaa et al., 2016; Lindström et al., 2010; Strömqvist et al., 2012).

Table 1. Hydrological and nitrogen-export processes, and parameter descriptions and ranges adopted in the HYPE model of the Selke catchment.

Processes	Parameters	Ranges
	Hydrological parameters	
Evapotranspiration	Evapotranspiration parameter *cevp* (land use dependent)	0.01-1.0 (mm/d/°C)
	Amplitude of sinus function that corrects potential evapotranspiration *cevpam* (general*)	0.01-1.0 (-)
	Phase of sinus function that corrects potential evapotranspiration *cevpph* (general*)	10-150 (d)
	Coefficient in exponential function for potential evapotranspiration's depth dependency *epotdist* (general*)	1-10 (1/m)
Surface flow and macro-pore flow	Recession coefficient for surface runoff *srrcs* (fraction, land use dependent)	0.01-1.0 (1/d)
	Fraction for surface runoff *srrate* (soil type dependent)	0.01-1.0 (-)
	Fraction for macro-pore flow *macrate* (soil type dependent)	0.01-1.0 (-)
	Threshold for macro-pore flow *mactrinf* (soil type dependent)	10-100 (mm/d)
	Threshold soil water for surface macro-pore flow and runoff *mactrsm* (fraction of wilting point + field capacity in uppermost layer, soil type dependent)	0.1-1.0 (-)
Soil interflow	Recession coefficient for uppermost soil layer *rrcs1* (soil type dependent)	0.01-1.0 (1/d)
	Recession coefficient for lowest soil layer *rrcs2* (soil type dependent)	0.0001-0.1 (1/d)
	Recession coefficient for slope dependence *rrcs3* (general*)	0.00001-0.001 (1/d/%)
Regional groundwater flow	Recession coefficient for regional groundwater outflow from soil layers *rcgrw* (general*)	0.0001-0.1 (-)
	Nitrogen parameters	
Nitrogen process in soil	Parameter for denitrification rate in soil *denitr* (general*)	0.001-0.1 (1/d)
	Decay of humusN to fastN *degradhn* (land use dependent)	0.00001-0.1 (1/d)
	Mineralization of fastN to inorganicN *minerfn* (land use dependent)	0.000001-0.1 (1/d)
	Fraction of nutrient uptake in the uppermost soil layer *uptsoil1* (land use dependent)	0.001-1.0 (-)
	Number of days that fertilizer applications occur *fertdays* (general*)	10-150 (d)
Nitrogen processes in stream	Production/decay of N in water *wprodn* (general*)	0.0001-0.1 ($kg/m^3/d$)
	Parameter for denitrification in water *denitw* (general*)	0.000001-0.1 ($kg/m^2/d$)
	Parameters for calculation of water velocity in watercourses *rivvel1, rivvel2, rivvel3*	0.01-1.0 (-)

* "general" means that the parameter is assumed to be applicable to the whole catchment.

2.3 Load estimation using different frequencies of nitrate measurement

The nitrate load was estimated from instantaneous flow and concentration measurements using:

$$L = \frac{K \sum_{i=1}^{n} (C_i Q_i)}{\sum_{i=1}^{n} Q_i} Q_r \tag{4}$$

$$Q_r = \frac{\sum_{j=1}^{N} Q_j}{N} \tag{5}$$

where L is the nitrate load for the period of interest (g/d), K converts time units (86,400 s/d), C_i is the instantaneous concentration (mg/L or g/m^3), Q_i is the instantaneous discharge at nutrient sampling time i (m^3/s), Q_r is the average discharge over the period of nitrate load estimation (m^3/s), Q_j is the recorded discharge at 15-min intervals (m^3/s), N is the number of flow measurements, and n is the number of concentration measurements.

2.4 Model setup

The HYPE model was set up to simulate a 5-year period, which includes calibration (i.e., 1st Nov 2010 to 31st Oct 2013) and validation (1st Nov 2013 to 31st Oct 2015), following the same procedure described by Jiang et al. (2014). The Selke catchment was divided into 29 sub-basins, ranging from 0.05 to 48 km^2, and averaging 15.1 km^2 in area. Nineteen soil types and ten land-use classes were adopted in categorizing the study area. Subsequently, 117 SLCs were defined by overlaying maps of soil type and land use. The hydrological driving data of daily rainfall and temperature were obtained from observations at precipitation and climate stations located within each sub-basin, or through interpolation to nearby stations. For streamflow simulation, hydrological parameters obtained from the multi-site calibration by Jiang et al. (2015), who used DREAM and data from the period 1994-1999, were used.

For nitrate modelling, the input data representing agricultural management (e.g., crop types, dates of sowing and harvest), fertilizer applications (rate and timing of applications of mineral fertilizer, organic fertilizer and manure), plant residue, and wet and dry atmospheric deposition of nitrogen were specified based on monitoring data, field surveys and literature review (e.g., Jiang et al., 2014; Kistner et al., 2013). Point source inputs, including daily discharge and average nitrate concentrations of wastewater outflows, were set based on the recordings from the six wastewater treatment plants within the Selke catchment.

2.5 Bayesian Inference Framework

Multi-site calibration was adopted in exploring posterior distributions of nitrogen-export process parameters. That is, nitrate datasets (at both daily and fortnightly frequencies) were obtained from all three gauging stations (Silberhuette, Meisdorf, and Hausneindorf) to better capture spatial variability in the nitrate export processes (Jiang et al., 2015). The calibration outputs using fortnightly and daily nitrate datasets were compared in terms of posterior parameter uncertainty and 95% prediction uncertainty of stream nitrate concentrations. Prior to calibration, sensitivity analysis was implemented using PEST to determine which parameters led to the largest modifications to model predictions. Parameter sensitivity is expressed as the Relative Composite Sensitivity (RCS), which measures the composite changes in model outputs incurred by a fractional change in the value of the parameter (Doherty, 2016; Jiang et al., 2014). The eight most sensitive nitrogen-export process parameters were selected for calibration, as listed in Table 2. A uniform prior distribution was assumed for each of the eight parameters, which were calibrated within the ranges specified in Table 2. Posterior parameter probability distributions and prediction uncertainty intervals were derived from calibration using the algorithm DREAM, which adopts the Bayesian inference framework (i.e., statistical inference, based on Bayes' theorem, of restrictions to the probability distribution of prior parameter distributions using informative field data and the likelihood function to yield posterior parameter distributions).

By applying Bayesian inference, the posterior parameter distributions can be derived by adjusting the parameters so that the spatiotemporal behavior of the model approximates, as closely and consistently as possible, the observed system behavior over some historical period of time (Balin et al., 2010). Two types of prediction uncertainty were assessed, namely parametric prediction uncertainty and total prediction uncertainty (Laloy and Vrugt, 2012). Bayesian inference treats the uncertainty of mechanistic model predictions in a very similar way to that of statistical models. Uncertainty is decomposed into that which can be explained by the uncertain model parameters (i.e., the parametric prediction uncertainty), and that which is attributable to other, uncalibrated factors. The uncertainty that cannot be explained by model parameters is characterised by a statistical distribution of residuals, i.e., the residual error. The residual error accounts for both measurement errors and model structural error, which are inseparable using the current approach, at least from a statistical perspective. Some insights into measurement error and model structural error are possible by undertaking multiple DREAM analyses using modified forms of the model, because by keeping the structure effectively the same, and by changing the number of measurements, it allows us to speculate on the likely breakdown of residual error into measurement error and model structural error. The parametric prediction uncertainty is computed by running HYPE on the posterior parameter distributions, following Balin et al. (2010). The total prediction uncertainty was determined using the statistical methods described below, and was taken to be the sum of the parametric prediction uncertainty and the residual error (Balin et al., 2010; Yang et al., 2008).

The method for assessing the uncertainty of parameters and predictions is obtained from Bayes' theorem, given by:

$$P(\theta|y^{obs}) = \frac{L(y^{obs}|\theta)P(\theta)}{\int L(y^{obs}|\theta)P(\theta)d\theta} \quad (6)$$

where $P(\theta)$ represents the assumed joint prior probability distribution of model parameters, contained in vector θ, $L(y^{obs}|\theta)$ is the likelihood function that quantifies the probability of measuring the field data y^{obs} given different θ values, and $P(\theta|y^{obs})$ is the posterior probability that expresses our updated beliefs in the θ values after the field data y^{obs} are taken into account through model calibration. The denominator is a scaling constant representing the sum of the conditional probabilities $L(y^{obs}|\theta)$ weighted by their prior probabilities $P(\theta)$ (Ellison, 1996). Sequences of model realizations from the posterior parameter distributions were obtained in our case using Markov chain Monte Carlo simulations.

The following likelihood function, $L(y^{obs}|\theta)$ was used to compare the simulated stream nitrate concentration dynamics with the observed nitrate datasets, assuming that the residuals between the observations and model results are identically, independently, and normally distributed (Yang et al., 2007):

$$L(y^{obs}|\theta) = \prod_{i=0}^{n} \left[\frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} exp\left(-\frac{1}{2} \frac{\left[y_{t_i}^{obs} - y_{t_i}^{M}(\theta) \right]^2}{\sigma^2} \right) \right] \quad (7)$$

where $y_{t_i}^{M}(\theta)$ and $y_{t_i}^{obs}$ represent, respectively, the simulated and observed nitrate concentrations at time t_i, and n is the total number of measurements. σ is the standard deviation of the differences between the simulated and observed nitrate concentrations, and is estimated jointly with model parameter θ.

Three Markov chains and a run length of 50,000 generations were used to ensure convergence of individual chains to the posterior distributions, following the suggested standard parameter setting of DREAM (Laloy and Vrugt, 2012). The posterior statistics were calculated using a thin of 10 (i.e., taking every 10^{th} sample instead of using the entire Markov chain, to minimize the effect of sample autocorrelation). We assessed convergence by visually inspecting plots of the posterior Markov chains for mixing and stationarity, and by inspecting density plots of the pooled posterior Markov chains for unimodality. We also assessed convergence quantitatively using the modified Gelman-Rubin convergence statistic (*R_stat*<1.2; Brooks and Gelman, 1998). The parameters and their physical meaning, and relevant statistical information, including prior uncertainty intervals, RCS and posterior statistics (maximum likelihood value (MAP), average value, and standard deviation (Std) of posterior parameter distributions) are listed in Table 2 for models based on daily and fortnightly nitrate datasets.

Table 2. Calibrated nitrogen-export process parameters and their prior and posterior statistics resulting from calibration to either fortnightly or daily nitrate datasets. Statistics include prior uncertainty intervals, Relative Composite Sensitivity (RCS), and posterior statistics (maximum likelihood value (MAP), average, and standard deviation (Std)).

Parameter	Physical meaning	Prior	Posterior statistics							
			Fortnightly nitrate dataset				Daily nitrate dataset			
			RCS	MAP	Average	Std	RCS	MAP	Average	Std
denitr	Parameter for denitrification rate in soil (1/d)	0.001-0.1	0.0043	0.029	0.035	0.0088	0.0086	0.027	0.026	0.0013
denitw	Parameter for denitrification in water ($kg/m^2/d$)	10^{-6}-0.1	5.6×10^{-7}	9.1×10^{-4}	7.6×10^{-4}	5.2×10^{-4}	1.8×10^{-6}	9.5×10^{-4}	9.7×10^{-4}	7.8×10^{-5}
wprodn	Production/decay of N in water ($kg/m^3/d$)	10^{-4}-0.1	1.6×10^{-7}	0.0048	0.0057	0.0035	1.7×10^{-6}	0.0040	0.0041	6.3×10^{-4}
uptsoil102	Fraction of nutrient uptake in the uppermost soil layer for arable land (-)	0.001-1	0.012	0.999	0.97	0.042	0.026	0.999	0.997	0.0036
uptsoil107	Fraction of nutrient uptake in the uppermost soil layer for coniferous forest (-)	0.001-1	0.0029	0.39	0.53	0.27	0.0052	0.0023	0.020	0.029
uptsoil108	Fraction of nutrient uptake in the uppermost soil layer for mixed forest (-)	0.001-1	6.3×10^{-5}	0.95	0.64	0.22	2.8×10^{-4}	0.96	0.96	0.010
rivvel2	Parameters for calculation of water velocity in watercourses (-)	0.01-1	1.1×10^{-9}	0.22	0.48	0.26	2.6×10^{-5}	0.22	0.22	0.0040
fertdays	Number of days that fertilizer applications occur (d)	10-150	0.0052	85	93	20	0.0089	78	78	1.9

2.6 Assessment of model performance and uncertainty

The model performance, in terms of streamflow, stream nitrate concentrations and nitrate loads, was evaluated using commonly adopted statistical criteria, including Nash-Sutcliffe coefficient (*NSE*), Percent BIAS (*PBIAS*), Average Absolute Error (*MAE*), and ratio of the root-average-square error to the standard deviation of field data (*RSR*) (Jiang et al., 2014; Laloy and Vrugt, 2012; Moriasi et al., 2007). Three criteria were used to evaluate the uncertainty of predictions of stream nitrate concentrations, including average relative interval length (*ARIL*), the percentage of measurements embodied by the 95% uncertainty interval (*PCI*), and the percentage of measurements bracketed by the unit uncertainty interval (*PUCI*). The concepts for these criteria are described in detail by Jin et al. (2010) and Li et al. (2011), and are therefore only summarised here. *ARIL* quantifies the sharpness of 95% uncertainty intervals of model predictions (i.e., parametric prediction uncertainty and total prediction uncertainty) and measures the resolution of the estimated prediction uncertainty. *PCI* assesses the reliability of the estimated prediction uncertainty. A smaller *ARIL* combined with a larger *PCI* represents narrower uncertainty intervals and higher reliability in the estimation of predictive uncertainty. *PUCI* is calculated using *ARIL* and *PCI*, as given below:

$$ARIL = \frac{1}{n}\sum \frac{Limit_{Upper,t} - Limit_{Lower,t}}{R_{obs,t}} \tag{8}$$

$$PUCI = (1.0 - Abs(PCI - 0.95))/ARIL \tag{9}$$

where $Limit_{Upper,t}$ and $Limit_{Lower,t}$ represent the respective upper and lower boundary values of the uncertainty interval of model predictions for the t^{th} day, n is the number of days, and $R_{obs,t}$ is the measured nitrate concentration on the t^{th} day. The larger the *PUCI*, the higher the confidence in the 95% prediction uncertainty interval limits as being representative of the model prediction uncertainty (i.e., both total prediction uncertainty and parametric prediction uncertainty were assessed using this method). The above criteria have proven to be informative in evaluating the uncertainty of model predictions in previous studies (Jin et al., 2010; Jiang et al., 2015; Li et al., 2011).

3. Results

3.1 Streamflow simulation

The streamflow dynamics evident in gauging station data are well represented by the HYPE model, with *NSE*≥0.70 (Figure 2, Table 3). The water balances are also relatively well captured, with the largest *PBIAS* of 11.6% reflecting underestimation at Hausneindorf during 1st Nov 2010 - 31st Oct 2013. Underestimation of peak flows in extraordinary storm events in January 2011 is noted (Figure 2). This can be explained in part by the uncertainty of rainfall data, whereby rainfall stations are sparsely distributed relative to changes in topography, which causes high spatial variability in rainfall (Woodward et al., 2017). Moreover, peak streamflow events are typically the result of short-term, intensive rainfall occurring at sub-daily time scales, and these are difficult to capture using a daily rainfall-runoff model (Jiang et al., 2014; Lam et al., 2012; Woodward et al., 2017).

There is a small decline of model performance during 1st Nov 2013 - 31st Oct 2015 relative to 1st Nov 2010 - 31st Oct 2013, as reflected by a decrease in *NSE*, and increases in *PBIAS* and *RSR*

(Table 3). This may be explained by considering that the hydrological conditions of these two periods are different, in that the later period is drier. A number of modelling studies have shown that hydrological parameters are not always temporally stable, and their values depend greatly on the length and physio-climatic conditions of the calibration period (e.g., Merz et al., 2009; Razavi and Tolson, 2013; Patil and Stieglitz, 2015). The relatively robust model performance in terms of streamflow for the entire simulation period indicates that the model structure and parameter set are sound and transferable between different meteorological (and therefore hydrological) conditions.

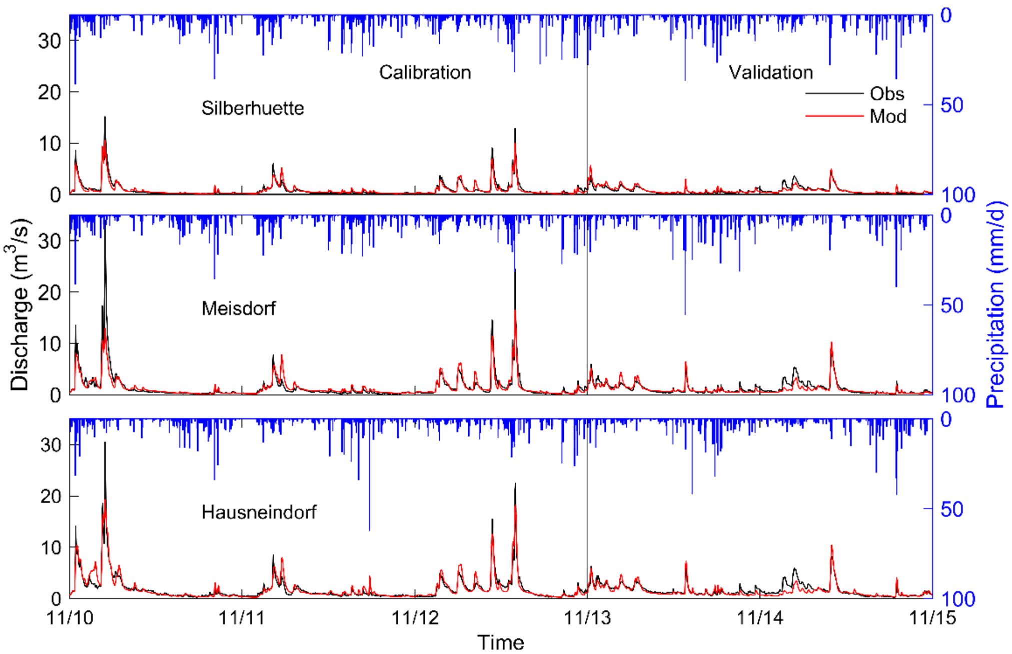

Figure 2. Streamflow simulation results together with observed daily rainfall during the periods of 1st Nov 2010 - 31st Oct 2013 (calibration) and 1st Nov 2013 - 31st Oct 2015 (validation) at Silberhuette, Meisdorf and Hausneindorf discharge gauging stations.

Table 3. Statistical model performance, in terms of streamflow at Silberhuette, Meisdorf and Hausneindorf gauging stations, during 1^{st} Nov 2010 - 31^{st} Oct 2013 and 1^{st} Nov 2013 - 31^{st} Oct 2015. *PBIAS* and *MAE* have the units of % and m^3/s, respectively, while *NSE* and *RSR* are unitless.

Sub-basin	1^{st} Nov 2010 - 31^{st} Oct 2013				1^{st} Nov 2013 - 31^{st} Oct 2015			
	NSE	*PBIAS*	*MAE*	*RSR*	*NSE*	*PBIAS*	*MAE*	*RSR*
Silberhuette	0.89	-2.9	0.26	0.33	0.70	6.5	0.23	0.55
Meisdorf	0.78	-2.5	0.49	0.47	0.74	-10.2	0.37	0.51
Hausneindorf	0.88	5.2	0.47	0.35	0.71	-11.6	0.42	0.54

3.2 Calibration of nitrogen-export process parameters

The sensitivity of nitrogen-export process parameters was tested by determining changes in simulated stream nitrate concentrations relative to changes in each parameter value. In application of PEST for sensitivity analysis, the overall sensitivity of each parameter was assessed by the magnitude of the respective vector of the Jacobian matrix, which contains derivatives of simulated nitrate concentrations with respect to parameter change. Based on sensitivity analysis using the daily nitrate dataset, the calibrated parameters are listed in order of most to least sensitive: *uptsoil102*, *fertdays*, *denitr*, *uptsoil107*, *uptsoil108*, *rivvel2*, *wprodn*, and *denitw*. This is a similar order to that derived from sensitivity analysis using the fortnightly nitrate dataset (parameters *rivvel2* and *denitw* are reversed in their order when the fortnightly dataset was adopted). As *uptsoil102* is the most sensitive parameter, this suggests that plant uptake in agricultural land is an important process controlling the nitrate balance within HYPE simulations. The parameters *fertdays* and *denitr*, which designate fertilizer applications and denitrification within the soil, respectively, are also very sensitive, as expected.

Except for the parameter *uptsoil107*, the maximum likelihood parameter values derived from calibration using fortnightly and daily datasets were similar. However, the posterior parameter distributions derived from calibration using fortnightly and daily nitrate datasets, quantified in terms of average parameter values and standard deviations (Table 2), were decidedly different. Specifically, the daily nitrate dataset produced much narrower posterior parameter uncertainty intervals relative to the fortnightly nitrate dataset, as indicated by standard deviations that were lower by more than an order of magnitude, except for *denitr* (Table 2). That is, it appears from these results that application of the daily dataset improved significantly the confidence in nitrogen-export process parameters.

3.3 Nitrate export simulation

Seasonal stream nitrate variations are characterized by high winter concentrations and low summer concentrations (Figure 3). The dynamics and overall nitrate balance were well captured at upland stations (Silberhuette and Meisdorf) following calibration to either daily or fortnightly nitrate datasets, with the lowest *NSE* equal to 0.43 and the largest *PBIAS* equal to 17.3% (Table 4). Nitrate concentrations in the stream remain high during high-flow conditions in winter. This is attributable to the following processes: (a) nitrate contained in soil and groundwater stores is flushed by higher rates of subsurface flow, (b) plant uptake is rather low (the average simulated uptake rate was lowest on average during January) due to low temperature and almost no agricultural activity. Conversely, nitrate concentrations in summer are low because (a) lower runoff leads to longer retention times within the catchment (stream discharge was lowest on average during August), which creates reduced rates of nutrient discharge to streams as nutrient storage levels build in the catchment, and subsurface nitrate losses occur prior to discharge to streams, and (b) higher temperatures and more intensive agricultural activities increase biogeochemical activity, leading to enhanced plant uptake (the monthly average uptake was highest in June). The dominant influence of subsurface flow on nitrate export and the strong effect of temperature on nitrate seasonal variability were also reported by Shrestha et al. (2007) from their nitrate export modelling study of Weida catchment (central Germany). The underestimation of nitrate concentrations in high-flow events at upland sites is perhaps caused partly by the underestimation of runoff (and the associated flushing of nutrients from the subsurface), although there may also be enhanced subsurface contributions to stream nitrate loads that are not well captured by the simple representation of aquifers in HYPE. At Hausneindorf, nitrate variability is lower than at the upland stations (Silberhuette, Meisdorf). A declining trend in nitrate concentrations was observed at the lowland station Hausneindorf, whereby the average measured nitrate concentration was 3.28 mg/L in the calibration period and 1.28 mg/L in the validation period. There is some evidence of overestimation of nitrate concentrations during low-flow conditions in summer that is probably attributable to the omission of denitrification in deep groundwater flow in the current model, and measurement errors/knowledge gaps related to nutrient sources.

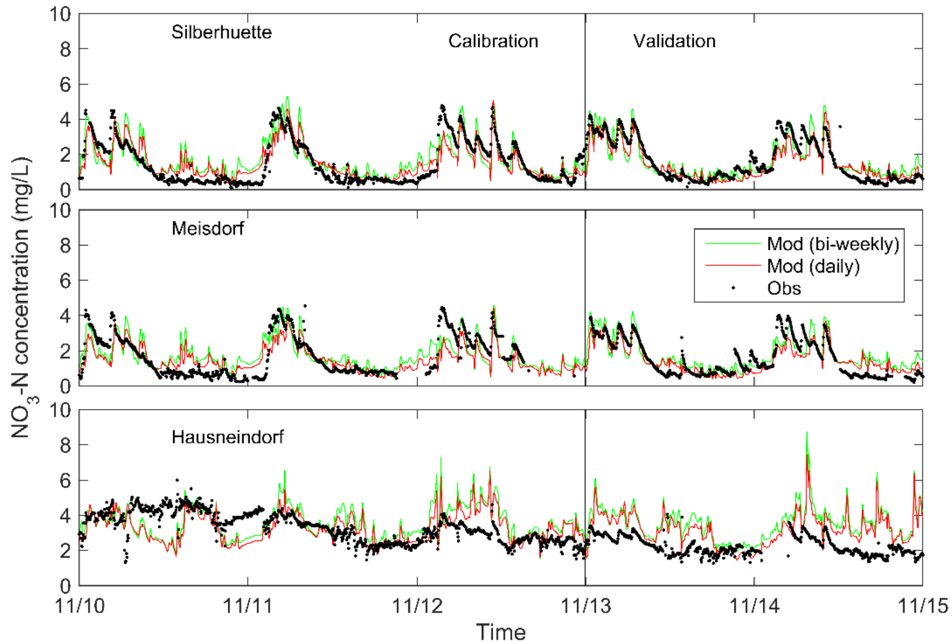

Figure 3. Observed (Obs) and Simulated (Mod) stream nitrate concentrations during calibration (1st Nov 2010 - 31st Oct 2013) and validation (1st Nov 2013 - 31st Oct 2015) periods. Results are shown for the three gauging stations (Silberhuette, Meisdorf and Hausneindorf), and as obtained from calibration using fortnightly and daily nitrate datasets.

Table 4. Statistical model performance in terms of stream nitrate concentrations and nitrate loads for calibration (1st Nov 2010 - 31st Oct 2013) and validation (1st Nov 2013 - 31st Oct 2015) periods. Results are given for the three gauging stations (Silberhuette, Meisdorf and Hausneindorf), and as obtained from calibration using fortnightly and daily nitrate datasets. The columns "Fortnightly" and "Daily" refer to the results of models calibrated to, respectively, fortnightly and daily nitrate datasets. The goodness-of-fit statistics refer to model-measurement comparisons based on daily nitrate datasets (i.e., daily average stream nitrate concentrations, and daily nitrate loads estimated as the product of streamflow and nitrate concentration).

Variable	Criterion	Calibration						Validation					
		Silberhuette		Meisdorf		Hausneindorf		Silberhuette		Meisdorf		Hausneindorf	
		Fortnightly	Daily	Fortnightly	Daily	Fortnightly	Daily	Fortnightly	Daily	Fortnightly	Daily	Fortnightly	Daily
Daily average nitrate concentr-ations	*NSE*	0.57	0.68	0.43	0.52	-0.95	-0.56	0.74	0.76	0.61	0.66	-8.55	-5.53
	PBIAS (%)	17.3	4.75	11.2	-4.73	7.81	-1.25	4.60	-8.02	15.1	-3.09	58.4	43.9
	MAE (mg/L)	0.63	0.52	0.69	0.62	0.92	0.79	0.47	0.43	0.53	0.46	1.32	1.07
	RSR	0.65	0.57	0.75	0.70	1.40	1.25	0.51	0.49	0.63	0.58	3.09	2.55
Daily nitrate loads	*NSE*	0.77	0.75	0.54	0.50	0.71	0.78	0.67	0.74	0.76	0.69	0.29	0.46
	PBIAS (%)	-8.56	-19.6	-19.3	-32.2	22.7	13.6	3.61	-10.2	-11.5	-25.4	48.8	34.9
	MAE (kg/d)	92.4	92.4	193	204	262	225	69.4	58.5	87.6	91.5	203	172
	RSR	0.48	0.50	0.68	0.71	0.54	0.47	0.58	0.51	0.49	0.55	0.84	0.74

At Silberhuette, model performance (in terms of nitrate concentrations) using parameters calibrated against the daily nitrate dataset is almost consistently better than that obtained from calibration against the fortnightly nitrate dataset, except for the slightly larger bias during the validation period (Table 4). At Meisdorf and Hausneindorf, the daily nitrate dataset outperformed the fortnightly nitrate dataset for both calibration and validation periods, as shown by the higher *NSE*, and lower *PBIAS*, *MAE* and *RSR*. This indicates that nitrogen-export process parameters calibrated using the daily nitrate dataset better represent nitrate dynamics occurring at a daily resolution relative to the parameters obtained from calibration to fortnightly nitrate sampling.

Whereas the abovementioned analysis of the daily and fortnightly datasets relied on comparison of daily predictions and measurements, the performance of both the daily-calibrated and fortnightly-calibrated models using fortnightly model outputs and measurements was also evaluated. This was aimed at testing whether the fortnightly-calibrated model better simulated fortnightly measurements relative to the daily-calibrated model – i.e., to evaluate whether the time-step of calibration controls which measurement frequency (daily or fortnightly) is best reproduced by the model. The results are given only in summary, for brevity. The results indicate that the daily-calibrated model better reproduces fortnightly measurements relative to the fortnightly-calibrated model. This is demonstrated by *NSE* values closer to one, and *PBIAS* values closer to zero for the daily-calibrated model for all stations. Specifically (values in brackets are for fortnightly-calibrated and daily-calibrated models, respectively): at Silberhuette, *NSE*=(0.56, 0.67) and *PBIAS*=(26.4%, 12.5%); at Meisdorf, *NSE*=(0.41, 0.47) and *PBIAS*=(18.6%, 0.34%), and at Hausneindorf, *NSE*= (-0.98, -0.50) and *PBIAS*=(15.0%, 5.0%).

The good reproduction of seasonal dynamics of streamflow and nitrate concentrations leads to daily nitrate loads that are reasonably well captured by the daily-calibrated model (Figure 4). The underestimation of daily nitrate loads during storm flow events is the combined result of the underestimation of streamflow and the underestimation of nitrate concentrations, as described earlier (see Figures 2 and 3). Daily nitrate loads are overestimated during low-flow conditions, especially at Hausneindorf, where the model overestimates nitrate concentrations, although the influence on annual nitrate loads is limited given the small contribution of low-flow conditions to annual loads. Calibration against "measured" nitrate loads (i.e., the product of measured streamflow and nitrate concentration) would likely improve load estimates obtained from HYPE. However, the focus here is on nitrate concentrations rather than loads, so we prefer not to increase errors in reproducing streamflow or nitrate concentrations by adding nitrate loads to the calibration objective function. In any case, HYPE captures the temporal and spatial variability of daily nitrate loads satisfactorily using parameters calibrated against either daily or fortnightly nitrate datasets, as demonstrated through the calibration and validation statistics given in Table 4.

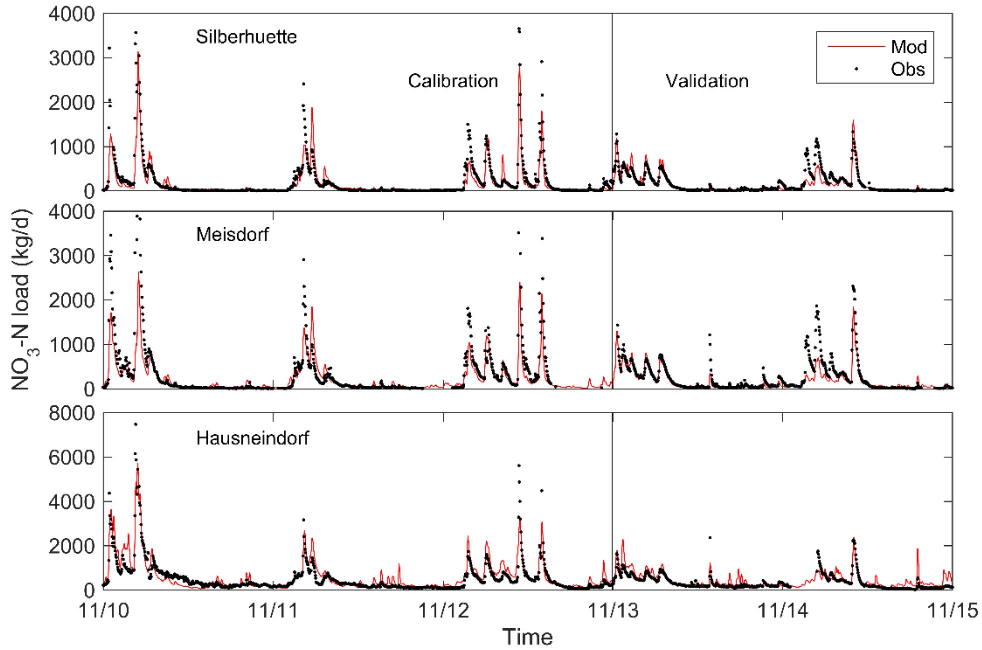

Figure 4. Observed (Obs) and simulated (Mod) daily nitrate loads during calibration (1st Nov 2010 - 31st Oct 2013) and validation (1st Nov 2013 - 31st Oct 2015) periods at the three gauging stations of Silberhuette, Meisdorf and Hausneindorf. "Observed daily nitrate loads" are nitrate loads calculated as the product of observed streamflow and nitrate concentration. "Simulated daily nitrate loads" represent the nitrate loads estimated using simulated streamflow and nitrate concentration from the model calibrated to daily data.

The spatial variations in predicted area-averaged nitrate loads (2010-2015) from calibration using fortnightly or daily nitrate data are similar, characterized by generally higher loads in lowland agricultural sub-basins and lower loads in upland forest sub-basins (Figure 5). Annual nitrate yields are not only controlled by land use, but are also highly dependent on hydrological regimes. For example, low nitrate loads are observed in some first-order lowland agricultural sub-basins. This is attributable to relatively smaller rates of runoff, which is a consequence of the shallow slope of the land surface and the lower rainfall of lowland areas. Also, subsurface fluxes of nutrients are reduced by the low soil permeability (chernozems). Conversely, relatively higher loads in upland forest sub-basins arise where the runoff is higher due to steeper slopes and where there is higher permeability in the upland soils (cambisols), leading to shorter retention times and enhanced transport of nitrate to streams. The variability in annual nitrate loads obtained from HYPE are within reasonable ranges compared to that obtained from nitrate export modelling studies of catchments with similar meteorological, hydrological and land use patterns. For example, our nitrate loads fall in the range of 0.15-21.99 kg/ha/yr, while Ahmad et al. (2011) and Rode et al. (2009) obtained ranges of a similar order, namely 0.84 to 4.98 kg/ha/yr and 18.5 to 41.2 kg/ha/yr, respectively, for catchments with comparable characteristics to the Selke catchment.

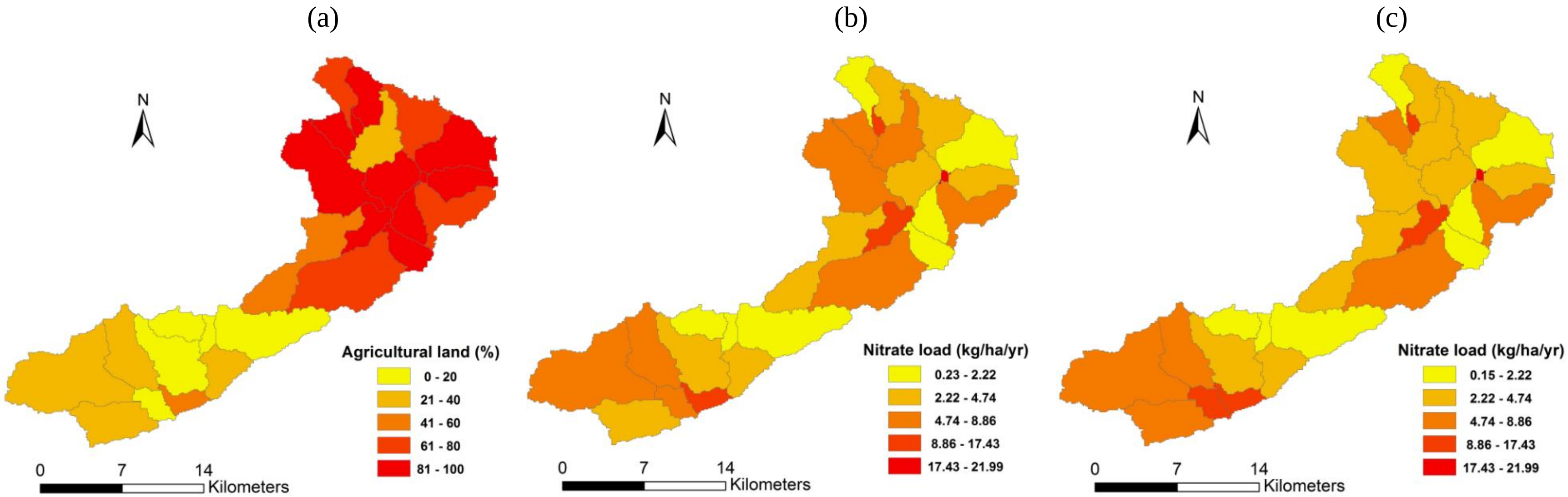

Figure 5. Simulated time- and area-averaged nitrate loads (kg/ha/yr) at Selke catchment during 1st Nov 2010 - 31st Oct 2015 following calibration using daily and fortnightly nitrate datasets. (a) percentage of agricultural land; (b) average nitrate loads following calibration against daily nitrate data; (c) average nitrate loads following calibration against fortnightly nitrate data.

3.4 Uncertainty in nitrate concentration predictions

The time-averaged parametric prediction uncertainty interval and total prediction uncertainty interval of nitrate concentrations over the calibration period at Silberhuette, obtained from calibration using the fortnightly dataset, are 1.40±0.23 mg/L and 1.40±1.95 mg/L, respectively. The values reflect the time-averaged predicted nitrate concentration plus-minus half of the difference between the upper and lower limits of time-averaged nitrate predictions. In comparison, the corresponding time-averaged uncertainty intervals obtained from calibration using the daily dataset are 1.49±0.056 mg/L and 1.49±1.80 mg/L, respectively. At Meisdorf, the time-averaged uncertainty intervals obtained from calibration using the daily dataset are 1.56±0.058 mg/L (parametric prediction uncertainty) and 1.56±1.80 mg/L (total prediction uncertainty), while corresponding values of 1.37±0.24 mg/L and 1.37±1.95 mg/L were obtained from calibration using the fortnightly dataset. At Hausneindorf, the daily dataset led to corresponding time-averaged uncertainty intervals of 3.26±0.090 mg/L and 3.26±1.80 mg/L, while 3.26±0.40 mg/L and 3.26±1.98 mg/L were obtained from calibration to fortnightly data.

The calibration results indicate that increasing the measurement frequency from fortnightly to daily also led to a four-fold reduction in the posterior parameter uncertainty (Table 2). However, in both the daily and fortnightly models, the parametric prediction uncertainty intervals are only a small proportion (3% (daily models) and 12% (fortnightly models) at both Silberhuette and Meisdorf, and 5% (daily models) and 20% (fortnightly models) at Hausneindorf) of the total prediction uncertainty intervals. The small contribution of parametric prediction uncertainty to total prediction uncertainty is also reflected by the much lower *ARIL* for both calibrated models using the alternative measurement frequencies (Figure 6 and Table 5).

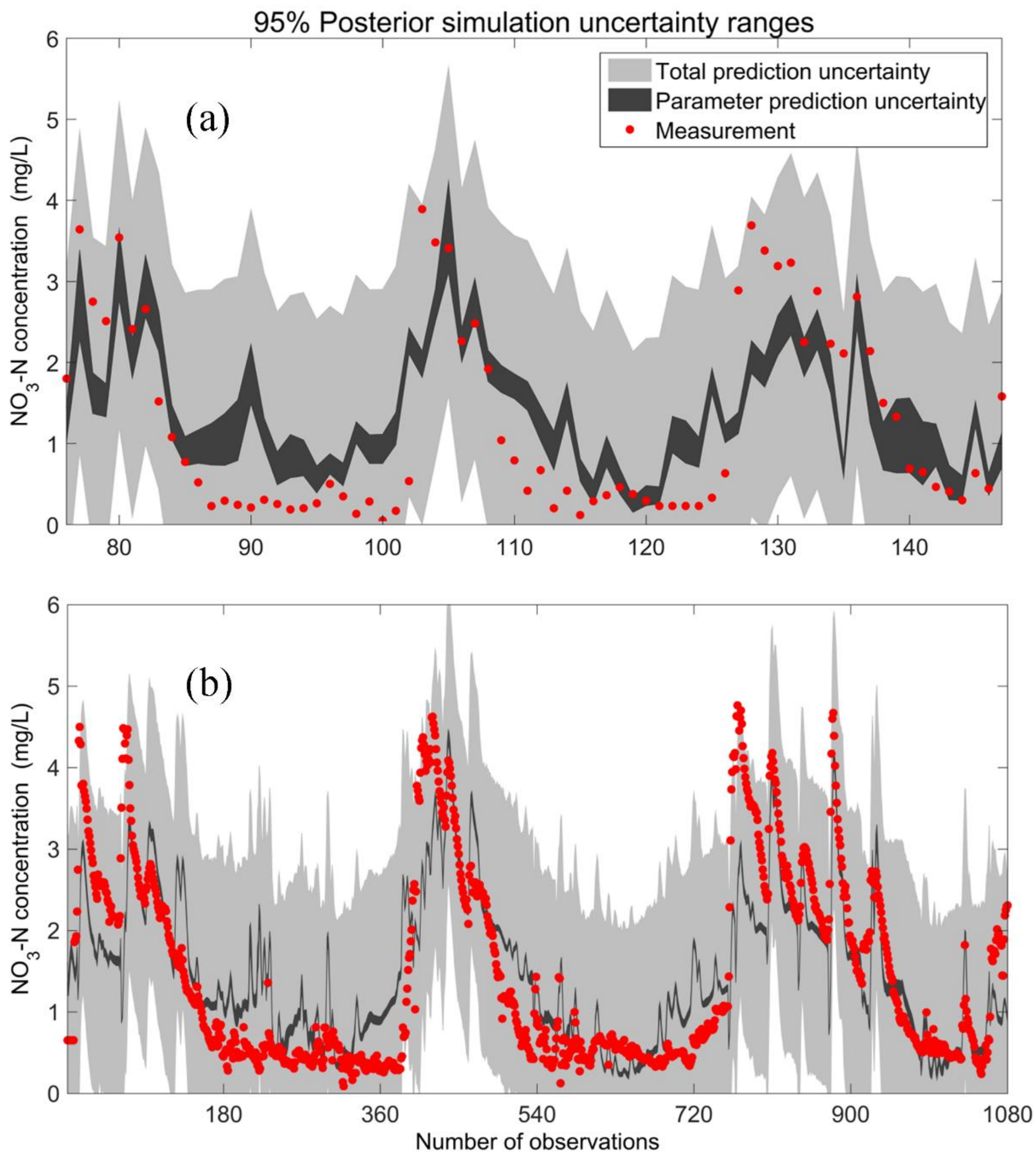

Figure 6. Comparison of 95% prediction uncertainty intervals of nitrate concentrations at Silberhuette during the period 1st Nov 2010 - 31st Oct 2013, estimated from calibration using: (a)

fortnightly, and (b) daily nitrate datasets. Black bands represent parametric prediction uncertainty intervals, grey bands represent total prediction uncertainty intervals, and red dots represent the corresponding stream nitrate measurements.

Table 5. Comparison of 95% prediction uncertainties in simulated stream nitrate concentrations at Silberhuette during 1st Nov 2010 - 31st Oct 2013, estimated from calibration using fortnightly and daily nitrate datasets. Par-Unc represents 95% parametric prediction uncertainty intervals of nitrate concentrations (mg/L); Tot-Unc represents 95% total prediction uncertainty intervals of nitrate concentrations (mg/L).

Criteria	Fortnightly nitrate dataset		Daily nitrate dataset	
	Par-Unc	Tot-Unc	Par-Unc	Tot-Unc
ARIL	0.97	9.2	0.15	4.7
PCI (%)	25	100	8.1	97
PUCI	0.31	0.10	0.87	0.21

Changing from fortnightly to daily calibration datasets produced only a modest lowering of the total prediction uncertainty, which was almost entirely attributable to the reduction in the posterior parameter uncertainty. Thus, the residual error was only slightly modified by the change to the calibration data frequency, with residual error in the model calibrated to daily data slight larger (9% at Silberhuette and Meisdorf, and 15% at Hausneindorf) than that of the model calibrated to fortnightly data. Inferences can be drawn from the finding that residual error is largely unmodified between the two calibration cases. For example, it is likely that the 14-fold increase (approximately) in calibration data between the fortnightly- to daily-calibrated models introduces larger measurement errors into the estimation of total prediction uncertainty, and an increase in residual error is a logical outcome of this. If the small increase in residual error is entirely the result of measurement error, then measurement errors are themselves rather small given the approximately 14-fold increase in the dataset size. This presumes that model structural error remains unchanged between the two calibration attempts. Given that the vast majority of the model design is identical in both calibration attempts, this presumption seems plausible. If we speculate based on these results that measurement error is indeed small, it infers that the model structural error is considerable, because it would consequently account for the majority of the total prediction uncertainty given that parametric prediction uncertainty is also modest.

As calibration using daily nitrate dataset results in decreased posterior parameter uncertainty (Table 2) and predictive uncertainty (Figure 6 and Table 5), we take from this that daily stream nitrate measurements enhance the confidence of predicted nitrate export. The assertions made here that total nitrate predictive uncertainty is mainly attributed to the combination of model structural error and measurement errors in our case is consistent with findings from previous studies (e.g., Balin et al., 2010; Jiang et al., 2015; Wellen et al., 2014; Woodward et al., 2017; Yang et al., 2007).

About 25% and 8.1% of nitrate measurements are contained within the 95% parametric prediction uncertainty intervals, as derived from calibration using fortnightly and daily nitrate datasets, respectively. The 95% total prediction uncertainty intervals capture 100% and 97.3% of nitrate measurements for models calibrated to fortnightly and daily nitrate datasets, respectively, indicating that both nitrate datasets produce reliable total prediction uncertainty intervals. The *PUCI* estimated from calibration to daily nitrate dataset was much larger than that obtained from calibration using the fortnightly nitrate dataset, in terms of both parametric prediction uncertainty intervals and total prediction uncertainty intervals (Table 5). This suggests that daily nitrate measurements are more informative to the calibration process, and subsequently offer greater constraints on nitrate transport and transformation parameters used in HYPE. Consequently, the daily nitrate dataset is more likely to produce unique parameters.

4. Discussion

The uncertainty analysis results highlight that, regardless of the frequency of calibration data, residual errors (model structural error and measurement error) dominate the uncertainty of the HYPE predictions of nitrate concentrations and loads. Calibration aims to reduce other uncertainty sources, and hence, the primary source of uncertainty within our modelling framework is immutable to calibration. Given that the alternative measurement frequencies considered in this study translate to model modifications via calibration, it seems probable that changing the measurement frequency is unlikely to affect model performance in a major way. Indeed, the greater number of measurements within the higher-frequency dataset may lead to increased measurement error, depending on several complicating factors. In our case, the residual error increased only slightly with the reduction in the measurement frequency. Woodward et al. (2017) obtained similar findings from their nutrient modelling study of the Weida catchment (Germany). They compared daily and monthly calibrations datasets spanning a four-year period, and found that monthly nitrate sampling provided sufficient information for the calibration of their low-spatial-resolution catchment model (StreamGEM). They evaluated sources of uncertainty using similar methods to those implemented here, and found, consistent with our results, that model structure and measurement error dominate total prediction uncertainty. However, they did not assess the impact of nitrate measurement frequencies on the posterior parameter uncertainty of their parsimonious catchment model. They were unable to separate measurement error and model structural error within the total prediction uncertainty, i.e., the same limitation of our method.

Despite the dominant role of residual error in the total prediction uncertainty, calibration using daily nitrate data reduced total prediction uncertainty, relative to fortnightly sampling (Table 2, Table 5 and Figure 6), albeit to a modest degree. While model reliability was only marginally better following calibration to the higher-frequency data, significant reductions in parametric prediction uncertainty were achieved. Thus, the key advantage of the higher frequency data was that higher confidence was instilled in the model parameters arising from the improved calibration of nitrogen-export process parameters. In practical terms, this means that the nitrate transport and transformation processes can be investigated with higher certainty, and the model is expected to provide useful assistance in refining the model to improve model structure and to guide the parameterization of nutrient inputs.

While residual error dominated total prediction uncertainty and inhibited the benefit of raising the frequency of calibration data in our case, we expect that the total predictive uncertainty of models with smaller residual error (lower model structural error and/or measurement error) will be reduced more significantly from higher-frequency water quality monitoring. There are several options for reducing the structural error in the current HYPE model. For instance, denitrification could be added to the deep groundwater flow to account for the nitrate losses and attenuation in nitrate inputs that occur within longer subsurface flow paths prior to the base flow discharge of groundwater to streams. Additional sampling of the groundwater chemistry and heads would also assist in understanding nitrate processes in the subsurface, and the groundwater impacts on stream nitrate concentrations more generally. Other advancements to the HYPE model are also possible for reducing structural error. For example, the simulation of in-stream biogeochemical processes could be modified to better account for the effects of flow velocity, sediment properties and light availability, etc. Higher resolution nitrate monitoring in stream networks enables the temporal and spatial variability in nitrate transformations in response to changes in climatic-hydrological conditions and biogeochemical activities to be captured. Higher resolution inputs to watershed nitrate export models allows for improved parameterization of in-stream biogeochemical processes and nutrient uptake and exchange rates. Additionally, the importance of nutrient retention in estimating nutrient loads at the catchment scale has been demonstrated in previous investigations (e.g., Grizzetti et al., 2003).

Our investigation would benefit from further efforts to match nutrient storage levels to field measurements of nutrient mass in the various hydrological components of the Selke catchment. The model's structure could also be modified to create greater temporal and spatial flexibility within the calibration process, such that parameters that are presently fixed in time and/or space are discretized to finer resolutions. For example, some parameters in HYPE (e.g., *uptsoil1*, *denitr*) that are presently constant in time could be at least seasonally variable. The benefits of adopting temporally variable, rather than time-invariant, parameters to account for variability in nutrient processes and hydrological controls has been demonstrated in previous rainfall-runoff and nutrient export investigations (e.g., Kim, 2016; Vandenberghe et al., 2007; Viswanathan et al., 2016). However, temporal variability is difficult to justify for some parameters, both from the perspective of physical process understanding and in terms of the availability of field data to constrain temporal parameter trends. Higher resolution nitrate data may allow for the estimation of temporally variable parameters (e.g., between low and high flow conditions), thereby reducing errors in process-based watershed nutrient export modeling that arise from the assumption of parameter time-invariance. This may address the problem that arises from lower frequency water quality monitoring, in that short-period flow events that produce the greatest nutrient fluxes are poorly resolved (Rode et al., 2016b).

The evaluation of model uncertainty has direct consequences for catchment management strategies that focus on nutrient transport. For example, the benefit of daily sampling depends on the availability of high-quality field monitoring that impart small measurement errors on the calibration effort. Knowledge of high-frequency nutrient behavior also allows for greater understanding of catchment processes, leading to model designs that better represent the key factors that impact nutrient concentrations and loads, and thereby lowering the model structural error. Also, the results of this investigation offer useful insights into catchment nutrient behavior (e.g., the spatial and temporal variability in nutrient fluxes, assessment of the nutrient balance for the study area, etc.) for the purposes of managing human activities of this region.

In order to develop guidance on water quality monitoring for the current study area, and in extending the findings to other regions, the results of this investigation need to incorporate other sources of information. For example, the design of water quality monitoring programs, in terms of selecting appropriate monitoring parameters, sampling sites, and sampling frequencies, is dependent on the monitoring objectives, budgetary constraints, and field-dependent characteristics (e.g., catchment size, land use, episodicity of rainfall, etc.; Strobl et al., 2006) that influence the need to capture short-term fluctuations and extremes. Indeed, according to the review of existing water quality monitoring strategies by Behmel et al. (2016), there is no holistic solution or guidance that covers all aspects of water quality monitoring programs. Nevertheless, the current study offers helpful insights to assist practitioners in selecting an appropriate temporal resolution of nitrate monitoring. For example, we show that fortnightly sampling of stream nitrate concentrations may be sufficient for credible calibration of nitrogen-export process parameters and satisfactory simulation of nitrate loads. Alternatively, daily sampling may be necessary to improve the confidence in calibration of nitrogen-export process parameters and to decrease parametric and overall prediction uncertainty of stream nitrate concentrations. While the choice of monitoring frequency will depend on several factors (e.g., measurement accuracy, modelling objectives, knowledge of nutrient transport processes, etc., as discussed above), in catchments similar to Selke, fortnightly sampling appears to be adequate for the calibration of semi-distributed models such as HYPE, if assessment of model uncertainty is not key to the aims of model development. Otherwise, daily monitoring data are preferred for attaining defensible and robust model parameters that more likely produce reasonable future predictions of catchment behavior.

In addition to the modeling outcomes, the methodology adopted in our investigation, in the form of combined application of DREAM and HYPE, provides a useful and novel example of uncertainty evaluation of a nitrate export model. We contend that the methodology adopted here offers a blueprint for investigating other catchments where the role of chemical measurement frequency, or other aspects of model development, are of interest. While there is an unavoidable element of localization in the uncertainty findings, it would be arguably less convincing to attempt to draw the same conclusions from a synthetic example or to use non-specific theoretical constructs in developing knowledge of the effects of sampling frequency on model uncertainty. The current research uses a real-world example to provide the first attempt at assessing in a comprehensive manner model uncertainty accompanying alternative temporal resolutions of sampling. We anticipate that others will report their findings regarding sampling frequency for different catchments and modeling methodologies over time, thereby building on our initial efforts to provide guidance on hydrochemical sampling frequency. Perhaps future research can attempt to assess nutrient behavior at the catchment scale based on calibration to sub-daily water quality monitoring. This may assist in advancing the current understanding of nutrient transport processes, although the current findings indicate that the application of sub-daily measurements would require lower model structural error. This is beyond the capability of the current model and the scope of this study. Other opportunities to extend the current study include the application of alternative uncertainty analysis strategies, aimed at identifying the relative contributions of individual modelling components to nitrate export uncertainty. This could assist in determining which model structures should be modified to create lower model structural error, to confirm (or otherwise) our conjecture that residual errors are most likely dominated by model structural errors rather than measurement errors within the total prediction uncertainty of the current modeling effort.

5. Conclusions

Fortnightly and daily stream nitrate measurements are compared in terms of their utility as calibration datasets in the simulation of nitrate export from a heterogeneous mesoscale catchment (Selke) in Germany. Comparisons are provided in terms of posterior parameter uncertainty of nitrogen-export process parameters, calibration mismatch, parametric prediction uncertainty, residual errors, and total prediction uncertainty. To this end, the DREAM code proved effective for the calibration and uncertainty analysis of a process-oriented catchment hydrological model (HYPE). Results show that calibrated nitrogen-export process parameter values using fortnightly and daily nitrate datasets are similar, although calibration using daily nitrate dataset generates much narrower posterior parameter uncertainty intervals. Thus, higher-resolution nutrient data led to greater confidence in model parameters.

The dynamics of daily stream nitrate concentrations and nitrate loads are captured satisfactorily at forest-dominated upland sub-basins without significant differences in model performance obtained from calibration using fortnightly and daily nitrate datasets. This suggests that fortnightly nitrate sampling provides sufficient measurements for calibration of nitrogen–export process parameters in regions of the model where agriculture is less intense, resulting in satisfactory simulation of nitrate export in forested catchments. Calibration using the daily nitrate dataset better represented fortnightly nitrate measurements relative to calibration using fortnightly nitrate sampling, and thus, the daily nitrate dataset has instilled better process representation in the model.

Mismatches in simulation of nitrate concentrations during low-flow conditions and at lowland agricultural sub-basin are noted. This is attributed primarily to model structural errors, such as (a) insufficient denitrification in deep groundwater flow, (b) over-simplification of in-stream biogeochemical processes, and (c) measurement errors in diffuse and point nutrient sources. The daily nitrate dataset produced significantly smaller parametric prediction uncertainty, but as this is only a small proportion of total prediction uncertainty, the higher frequency dataset led to only modest reduction in total prediction uncertainty.

This study concludes that changing nitrate measurement frequency did not have a significant effect on the reproduction of observed stream nitrate dynamics and the total uncertainty of nitrate predictions more generally, because the combination of model structural error and measurement errors were much higher relative to parametric prediction uncertainty. However, increasing measurement frequency could more significantly affect the accuracy of nitrate export simulation if the measurement errors are reduced, and more advanced model structures are developed and utilized in the future.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (41877487, 41501531), Natural Science Foundation of Jiangsu Province (BK20151062), and the Open Research Fund of State Key Laboratory of Simulation and Regulation of Water Cycle in River

Basin (China Institute of Water Resources and Hydropower Research, Grant NO: IWHR-SKL-201710). Special thanks go to Jasper Vrugt for sharing program codes of DREAM.

References

Ajami, N. K., Duan, Q. Y., and S. Sorooshian (2007), An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resour. Res.*, 43, Art. no. W01403.

Abbaspour, K. C., Johnson, A., and M. Th. van Genuchten (2004), Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone J*., 3, 1340-1352.

Ahmad, H. M. N., Sinclair, A., Jamieson, R., Madani, A., Hebb, D., Havard, P., and E. K. Yiridoe (2011), Modeling sediment and nitrogen export from a rural watershed in eastern Canada using the soil and water assessment tool. *J. Environ. Qual.*, 40, 1182-1194.

Arhonditsis, G. B., Qian, S. S., Stow, C. A., Lamon, E. C., and K. H. Reckhow (2007), Eutrophication risk assessment using Bayesian calibration of process-based models: Application to a mesotrophic lake. *Ecol. Modell.*, 208 (2-4), 215-229.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., and J. R. Williams (1998), Large area hydrological modeling and assessment part I: model development. *J. Am. Water Resour. Assoc.*, 34 (1), 73-89.

Aubert, A. H., Gascuel-Odoux, C., Gruau, G., Akkal, N., Faucheux, M., Fauvel, Y., Grimaldi, C., Hamon, Y., Jaffrézic, A., Lecoz-Boutnik, M., Molénat, J., Petitjean, P., Ruiz, L., and P. Merot (2013), Solute transport dynamics in small, shallow groundwater-dominated agricultural catchments: insights from a high-frequency, multisolute 10 yr-long monitoring study. *Hydrol. Earth Syst. Sci.*, 17 (4), 1379-1391.

Balin, D., Lee, H., and M. Rode (2010), Is point uncertain rainfall likely to have a great impact on distributed complex hydrological modeling? *Water Resour. Res.*, 46, Art. no. W11520.

Basu, N. B., Thompson, S. E., and P. S. C. Rao (2011), Hydrologic and biogeochemical functioning of intensively managed catchments: a synthesis of topdown analyses. *Water Resour. Res.*, 47, Art. no. W00J15.

Behmel, S., Damour, M., Ludwig, R., and M. J., Rodriguez (2016), Water quality monitoring strategies - A review and future perspectives. *Sci. Total Environ.*, 571, 1312-1329.

Beven, K., and A. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction. *Hydrol. Processes*, 6, 279-298.

Bicknell, B. R., Imhoff, J. C., Kittle, J. L. Jr., Jobes, T. H., and A. S. Donigian Jr. (2012), HSPF, Version 12, User's Manual. U.S. Environmental Protection Agency, Athens, GA.

Bingner, R. L., Theurer, F. D., and Y. Yuan (2012), AnnAGNPS Technical Processes. http://www.ars.usda.gov/Research/docs.htm?docid=5199 (access on 5/16/2017).

Brooks, S. P., and A. Gelman (1998), General methods for monitoring convergence of iterative Simulations. *J. Comput. Graph. Stat.*, 7, 434-455.

Chappell, N. A., Jones, T. D., and W. Tych (2017), Sampling frequency for water quality variables in streams: Systems analysis to quantify minimum monitoring rates. *Water Res.*, 123, 49-57.

Cao, W., Bowden, W. B., Davie, T., and A. Fenemor (2006), Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrol. Process.*, 20 (5), 1057-1073.

Conley, D. J., Paerl, H. W., Howarth, R. W., Boesch, D. F., Seitzinger, S. P., Havens, K. E., Lancelot, C., and G. E. Likens (2009), ECOLOGY controlling eutrophication: nitrogen and phosphorus. *Science*, 323, 1014-1015.

Doherty, J. (2016), PEST User Manual. Part I and II. Watermark Numerical Computing, Brisbane, Australia.

Dotto, C. B. S., Mannina, G., Kleidorfer, M., Vezzaro, L., Henrichs, M., McCarthy, D. T., Freni, G., Rauch, W., and A. Deletic (2012), Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling. *Water Res.*, 46 (8), 2545-2558.

Ellison, A. M. (1996), An Introduction to Bayesian Inference for Ecological Research and Environmental Decision-Making. *Ecol. Appl.*, 6 (4), 1036-1046.

Fovet, O., Ruiz, L., Faucheux, M., Molénat, J., Sekhar, M., Vertès, F., Aquilina, L., Gascuel-Odoux, C., and P. Durand (2015), Using long time series of agricultural-derived nitrates for estimating catchment transit times. *J. Hydrol.*, 522, 603-617.

Gong, Y. W., Shen, Z. Y., Hong, Q., Liu, R. M., and Q. Liao (2011), Parameter uncertainty analysis in watershed total phosphorus modeling using the GLUE methodology. *Agric. Ecosyst. Environ.*, 142, 246-255.

Grizzetti, B., Bouraoui, F., Granlund, K., Rekolainen, S., and G. Bidoglio (2003), Modelling diffuse emission and retention of nutrients in the Vantaanjoki watershed (Finland) using the SWAT model. *Ecol. Modell.*, 169, 25-38.

Halliday, S., Skeffington, R. A., Wade, A. J., Bowes, M. J., Gozzard, E., Newman, J. R., Loewenthal, M., Palmer-Felgate, E. J., and H. P. Jarvie (2015), High-frequency water quality monitoring in an urban catchment: hydrochemical dynamics, primary production and implications for the Water Framework Directive. *Hydrol. Process.*, 29, 3388-3407.

Horowitz, A. (2003), An evaluation of sediment rating curves for estimating suspendend sediment concentrations for subsequent flux calculations. *Hydrol. Process.*, 17, 3387-3409.

Hrachowitz, M., Benettin, P., van Breukelen, B. M., Fovet, O., Howden, N. J. K., Ruiz, L., van der Velde, Y., and A. J. Wade (2016), Transit times-the link between hydrology and water quality at the catchment scale. *WIREs Water*, doi: 10.1002/wat2.1155.

Jiang, S. Y., Jomaa, S., Buettner, O., Meon, O., and M. Rode (2015), Multi-site identification of a distributed hydrological nitrogen model using Bayesian uncertainty analysis. *J. Hydrol.*, 529, 940-950.

Jiang, S. Y., Jomaa, S., and M. Rode (2014), Modelling inorganic nitrogen emissions at a nested mesoscale catchment in central Germany. *Ecohydrology*, 7 (5), 1345-1362.

Jin, X., Xu, C. -Y., Zhang, Q., and V. P. Singh (2010), Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. *J. Hydrol.*, 383 (3-4), 147-155.

Jomaa, S., I. Aboud, R. Dupas, X. Yang, J. Rozemeijer, and M. Rode (2018), Improving nitrate load estimates in an agricultural catchment using Event Response Reconstruction, *Environ. Monit. Assess.*, 190 (6), 330.

Jomaa, S., Jiang, S. Y., Thraen, D., and M. Rode (2016), Modelling the effect of different agricultural practices on stream nitrogen load in central Germany. *Energy, Sustain. Soc.*, 6 (11), 1-16.

Jones, A. S., Horsburgh, J., Mesner, N. O., Ryel, R. J., and D. K. Stevens (2012), Influence of sampling frequency on estimation of annual total phosphorus and total suspended solids loads. *J. Am. Water Resour. Assoc.*, 48 (6), 1258-1275.

Jones, T. D. and N. A. Chappell (2014), Streamflow and hydrogen ion interrelationships identified using data-based mechanistic modelling of high frequency observations through contiguous storms. *Hydrol. Res.*, 45 (6), 868-892.

Jones, T. D., Chappell, N. A., and W. Tych (2014), First dynamic model of dissolved organic carbon derived directly from high-frequency observations through contiguous storms. *Environ. Sci, Technol.*, 48 (22), 13289-13297.

Jordan, P., and R. Cassidy (2011), Technical Note: Assessing a 24/7 solution for monitoring water quality loads in small river catchments. *Hydrol. Earth Syst. Sci.*, 15, 3093–3100.

Kim, H. S. (2016), Potential Improvement of the Parameter Identifiability in Ungauged Catchments. *Water Resour. Manage.*, 30 (9), 3207-3228.

Kim, D. -K., Kaluskar, S., Mugalingam, S., Blukacz-Richards, A., Long, T., Morley, A., and G. B. Arhonditsis (2017), A Bayesian apporach for estimating phosphorus export and delivery rates with the SPAtially Referenced Regression On Watershed attributes (SPARROW) model. *Ecol. Inf.*, 37, 77-91.

Kirchner, J. W., Feng, X. H., Neal, C., and A. J. Robson (2004), The fine structure of water-quality dynamics: the(high-frequency) wave of the future. *Hydrol. Processes*, 18 (7), 1353-1359.

Kistner, I., Ollesch, G., Meissner, R., and M. Rode (2013), Spatial-temporal dynamics of water soluble phosphorus in the topsoil of a low mountain range catchment. *Agric. Ecosyst. Environ.*, 48, 24-38.

Knowling, M. J., and A. D. Werner (2016), Estimability of recharge through groundwater model calibration: Insights from a field-scale steady-state example. *J. Hydrol.*, 540, 973-987.

Kyllmar, K., Bechmann, M., Deelstra, J., Iital, A., Blicher-Mathiesen, G., Jansons, V., Koskiaho, J., and A. Povilaitis (2014), Long-term monitoring of nutrient losses from agricultural catchments in the Nordic-Baltic region – A discussion of methods, uncertainties and future needs. *Agric. Ecosyst. Environ.*, 198, 4-12.

Laloy, E., and J. A. Vrugt (2012), High-dimensional posterior exploration of hydrologic models using multiple-try $\text{DREAM}_{(ZS)}$ and high-performance computing. *Water Resour. Res.*, 48, Art. no. W01526.

Lam, Q. D., Schmalz, B., and N. Fohrer (2012), Assessing the spatial and temporal variations of water quality in lowland areas, Northern Germany. *J. Hydrol.*, 438-439, 137-147.

Levine, C. R., Yanai, R. D., Lampman, G. G., Burns, D. A., Driscoll, C. T., Lawrence, G. B., Lynch, J. A., and N. Schoch (2014), Evaluating the efficiency of environmental monitoring programs. *Ecol. Indic.*, 39, 94-101.

Li, H., Sivapalan, M., Tian, F., and D. Liu (2010), Water and nutrient balances in a large tile-drained agricultural catchment: a distributed modeling study. *Hydrol. Earth Syst. Sci.*, 14, 2259-2275.

Li, L., Xu, C. -Y., Xia, J., Engeland, K., and P. Reggiani (2011), Uncertainty estimates by Bayesian method with likelihood of AR (1) plus Normal model and AR (1) plus Multi-Normal model in different time-scales hydrological models. *J. Hydrol.*, 406 (1-2), 54-65.

Liu, Y., and H. V. Gupta (2007), Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resour. Res.*, 43, Art. no. W07401.

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., and B. Arheimer (2010), Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrol. Res.*, 41 (3-4), 295-319.

Lloyd, C. E. M., Freer, J. E., Johnes, P. J., and A. L. Collins (2016), Using hysteresis analysis of high-resolution water quality monitoring data, including uncertainty, to infer controls on nutrient and sediment transfer in catchments. *Sci. Total Environ.*, 543, 388-404.

Merz, R., Parajka, J., and G. Bloechl (2009), Scale effects in conceptual hydrological modeling. *Water Resour. Res*., 45 (9), W09405.

Molenat, J., Gascuel-Odoux, C., Gascuel-Odoux, C. Ruiz, L., and G. Gruau (2008), Role of water table dynamics on stream nitrate export and concentration in agricultural headwater catchment (France). *J. Hydrol.*, 348 (3-4), 363-378.

Moriasi, D. N., Arnold, J. G., Liew, M. W. V., Bingner, R. L., Harmel, R. D., and T. L. Veith (2007), Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE*, 50, 885-900.

Onderka, M., Wrede, S., Rodný, M., Pfister, L., Hoffmann, L., and A. Krein (2012), Hydrogeologic and landscape controls of dissolved inorganic nitrogen (DIN) and dissolved silica (DSi) fluxes in heterogeneous catchments. *J. Hydrol.*, 450-451, 36-47.

Pathak, D., Whitehead, P. G., Futter, M. N., and R. Sinha (2018), Water quality assessment and catchment-scale nutrient flux modeling in the Ramganga River Basin in north India: An application of INCA model. *Sci. Total Environ.*, 631-632, 201-215.

Patil, S. D., and M. Stieglitz (2015), Comparing spatial and temporal transferability of hydrological model parameters. *J. Hydrol*., 525, 409-417.

Razavi, S., and B. A.Tolson (2013), An efficient framework for hydrologic model calibration on long data periods. *Water Resour. Res.*, 49 (12), 8418-8431.

Rode, M., Arhonditsis, G., Balin, D., Kebede, T., Krysanova, V., van Griensven, A., and S. E. A. T. M. van der Zee (2010), New challenges in integrated water quality modelling. *Hydrol. Processes*, 24, 3447-3461.

Rode, M., Halbedel, S., Anis, M. R., Borchardt, D., and M. Weitere (2016a), Continuous in-stream assimilatory nitrate uptake from high-frequency sensor measurements. *Environ. Sci. Technol.*, 50 (11), 5685-5694.

Rode, M., Thiel, E., Franko, U., Wenk, G., and F. Hesser (2009), Impact of selected agricultural management options on the reduction of nitrogen loads in three representative meso scale catchments in Central Germany. *Sci. Total Environ.*, 407, 3459-3472.

Rode, M., Wade, A. J., Cohen, M. J., Hensley, R. T., Bowes, M. J., Kirchner, J. W., Arhonditsis, G. B., Jordan, P., Kronvang, B., Halliday, S. J., Skeffington, R. A., Rozemeijer, J. C., Aubert, A. H., Rinke, K., and S. Jomaa (2016b), Sensors in the Stream: The High-Frequency Wave of the Present. *Environ. Sci. Technol.*, 50 (11), 10297-10307.

Rodríguez-Blanco, M. L., Taboada-Castro M. M., and M. T. Taboada-Castro (2013), Phosphorus transport into a stream draining from a mixed land use catchment in Galicia (NW Spain): Significance of runoff events. *J. Hydrol.*, 481, 12-21.

Ross, C., Petzold, H., Penner, A., and G. Ali (2015), Comparison of sampling strategies for monitoring water quality in mesoscale Canadian Prairie watershed. *Environ. Monit. Assess.*, 395.

Sandford, R. C., Hawkins, J. M., Bol, R., and P. J. Worsfold (2013), Export of dissolved organic carbon and nitrate from grassland in winter using high temporal resolution, in situ UV sensing. *Sci. Total Environ*., 456-457, 384-391.

Sharpley, A. N., Kleinman, P. J., Heathwaite, A. L., Gburek, W. J., Folmar G. J., and J. P. Schmidt (2008), Phosphorus loss from an agricultural watershed as a function of storm size. *J. Environ. Qual.*, 37, 362-368.

Shrestha, R. R., Bárdossy A., and M. Rode (2007), A hybrid deterministic–fuzzy rule based model for catchment scale nitrate dynamics. *J. Hydrol.*, 342, 143-156.

Shrestha, R. R., Dibike, Y. B., and T. D. Prowse (2012), Modeling Climate Change impacts on hydrology and nutrient loading in the upper Assiniboine catchment. *J. Am. Water Resour. Assoc.*, 48, 74-89.

Skarbøvik, E., Stålnacke, P., Bogen, J., and T. Bønsnes (2012), Impact of sampling frequency on mean concentrations and estimated loads of suspended sediment in a Norwegian river: Implications for water management. *Sci. Total Environ.*, 433, 462-471.

Strobl, R. O., Robillard, P. D., Shannon, R. D., Day, R. L., and A. J. McDonnell (2006), A water quality monitoring network design methodology for the selection of critical sampling points: Part I. *Environ. Monit. Assess.*, 112 (1-3), 137-158.

Strömqvist, J., Arheimer, B., Dahné, J., Donnelly, C., and G. Lindström (2012), Water and nutrient predictions in ungauged basins: Set-up and evaluation of a model at the national scale. *Hydrol. Sci. J.*, 57, 229-247.

Tonderski, K., Andersson, L., Lindström, G., Cyr, R. S., Schönberg, R., and H. Taubald (2017), Assessing the use of $\delta^{18}O$ in phosphate as a tracer for catchment phosphorus sources. *Sci. Total Environ.*, 607-608, 1-10.

Ullrich, A., and M. Volk (2010), Influence of different nitrate-N monitoring strategies on load estimation as a base for model calibration and evaluation. *Environ. Monit. Assess.*, 171, 513-527.

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., and R. Srinivasan (2006), A global sensitivity analysis tool for the parameters of multi-variable catchment models. *J. Hydrol.*, 324, 10-23.

Van Meter, K. J., and N. B. Basu (2015), Catchment Legacies and Time Lags: A Parsimonious Watershed Model to Predict the Effects of Legacy Storage on Nitrogen Export. *PLoS ONE*, 10 (5), e0125971.

Van Meter, K. J., Basu, N. B., Veenstra, J. J., and C. L. Burras (2016), The nitrogen legacy: emerging evidence of nitrogen accumulation in anthropogenic landscapes. *Environ. Res. Lett.*, 11, 035014.

Vandenberghe, V., Bauwens, W., and P. A. Vanrolleghem (2007), Evaluation of uncertainty propagation into river water quality predictions to guide future monitoring campaigns. *Environ. Modell. Software*, 22, 725-732.

Viswanathan, V. C., Jiang, Y., Berg, M., Hunkeler, D., and M. Schirmer (2016), An integrated spatial snap-shot monitoring method for identifying seasonal changes and spatial changes in surface water quality. *J. Hydrol.*, 539, 567-576.

Vrugt, J. A. (2016), Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. *Environ. Modell. Software*, 75, 273-316.

Vrugt, J. A., Gupta, H. V., Bouten, W., and S. Sorooshian (2003), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour. Res.*, 39 (8), 1-16.

Wade, A. J., Palmer-Felgate, E. J., Halliday, S. J., Skeffington, R. A., Loewenthal, M., Jarvie, H. P., Bowes, M. J., Greenway, G. M., Haswell, S. J., Bell, I. M., Joly, E., Fallatah, A., Neal, C., Williams, R. J., Gozzard, E., and J. R. Newman (2012), Hydrochemical processes in lowland rivers: insights from in situ, high-resolution monitoring. *Hydrol. Earth Syst. Sci.*, 16 (11), 4323-4342.

Wellen, C., Arhonditsis, G. B., Long,T., and D. Boyd (2014), Quantifying the uncertainty of nonpoint source attribution in distributed water quality models: A Bayesian assessment of SWAT's sediment export predictions. *J. Hydrol.*, 519, 3353-3368.

Wellen, C., Kamran-Disfani, A. -R., and G. B. Arhonditsis (2015), Evaluation of the current state of distributed nutrient watershed-water quality modeling. *Environ. Sci. Technol.*, 49 (6), 3278-3290.

Whitehead, P., Wilson, E., and D. Butterfield (1998), A semi-distributed ntegrated nitrogen model for multiple source assessment in tchments (INCA): Part I –model structure and process equations. *Sci. Total Environ.*, 210-211, 547-558.

Wollschläger, U., Attinger, S., Borchardt, D., Brauns, M., Cuntz, M., Dietrich, P., Fleckenstein, J. H., Friese, K., Friesen, J., Harpke, A., Hildebrandt, A., Jäckel, G., Kamjunke, N., Knöller, K., Kögler, S., Kolditz, O., Krieg, R., Kumar, R., Lausch, A., Liess, M., Marx, A., Merz, R., Mueller, C., Musolff, A., Norf, H., Oswald, S. E., Rebmann, C., Reinstorf, F., Rode, M., Rink, K., Rinke, K., Samaniego, L., Vieweg, M., Vogel, H.-J., Weitere, M., Werban, U., Zink, M., and S. Zacharias (2017), The Bode hydrological observatory: a platform for integrated, interdisciplinary hydro-ecological research within the TERENO Harz/Central German Lowland Observatory. *Environ. Earth Sci.*, 76 (1), art. 29.

Woodward, S. J. R., Wöhling, T., Rode, M., and S. Roland (2017), Predicting nitrate discharge dynamics in mesoscale catchments using the lumped StreamGEM model and Bayesian parameter inference. *J. Hydrol.*, 552, 684-703.

Wu, H. J., and B. Chen (2015), Evaluating uncertainty estimates in distributed hydrological modeling for the Wenjing River watershed in China by GLUE, SUFI-2, and ParaSol method. *Ecol. Eng*., 76, 110-121.

Xia, Y. Q., Weller, D. E., Williams, D. E., Jordan, T. E., and X. Y. Yan (2016), Using Bayesian hiterarchical models to better understand nitrate sources and sinks in agricultural watersheds. *Water Res*., 105, 527-539.

Yang, J., Reichert, P., Abbaspour, K. C., Xia, J., and H. Yang (2008), Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *J. Hydrol.*, 358 (1-2), 1-23.

Yang, J., Reichert, P., and K. C. Abbaspour (2007), Bayesian uncertainty analysis in distributed hydrologic modeling: A case study in the Thur River basin (Switzerland). *Water Resour. Res.*, 43, Art. no. W10401.

Yang, X., Jomaa, S., Zink, M., Fleckenstein, J. H., Borchardt, D., and M. Rode (2018), A new fully distributed model of nitrate transport and removal at catchment scale. *Water Resour. Res.*, Accepted Article, doi: 10.1029/2017WR022380.

Yin, Y. X., Jiang, S. Y., Pers, C., Yang, X. Y., Liu, Q., Jin, Y., Yao, M. X., He, Y., Luo, X. Z., and Z. Zheng (2016), Assessment of the Spatial and Temporal Variations of Water Quality for Agricultural Lands with Crop Rotation in China by Using a HYPE model. *Int. J. Environ. Res. Public Health*, 13 (3), 336, 1-19.

Table 1
Click here to download Table: Table 1.docx

Table 1. Hydrological and nitrogen-export processes, and parameter descriptions and ranges adopted in the HYPE model of the Selke catchment.

Processes	Parameters	Ranges
	Hydrological parameters	
Evapotranspiration	Evapotranspiration parameter *cevp* (land use dependent)	0.01-1.0 (mm/d/°C)
	Amplitude of sinus function that corrects potential evapotranspiration *cevpam* (general*)	0.01-1.0 (-)
	Phase of sinus function that corrects potential evapotranspiration *cevpph* (general*)	10-150 (d)
	Coefficient in exponential function for potential evapotranspiration's depth dependency *epotdist* (general*)	1-10 (1/m)
Surface flow and macro-pore flow	Recession coefficient for surface runoff *srrcs* (fraction, land use dependent)	0.01-1.0 (1/d)
	Fraction for surface runoff *srrate* (soil type dependent)	0.01-1.0 (-)
	Fraction for macro-pore flow *macrate* (soil type dependent)	0.01-1.0 (-)
	Threshold for macro-pore flow *mactrinf* (soil type dependent)	10-100 (mm/d)
	Threshold soil water for surface macro-pore flow and runoff *mactrsm* (fraction of wilting point + field capacity in uppermost layer, soil type dependent)	0.1-1.0 (-)
Soil interflow	Recession coefficient for uppermost soil layer *rrcs1* (soil type dependent)	0.01-1.0 (1/d)
	Recession coefficient for lowest soil layer *rrcs2* (soil type dependent)	0.0001-0.1 (1/d)
	Recession coefficient for slope dependence *rrcs3* (general*)	0.00001-0.001 (1/d/%)
Regional groundwater flow	Recession coefficient for regional groundwater outflow from soil layers *rcgrw* (general*)	0.0001-0.1 (-)
	Nitrogen parameters	
Nitrogen process in soil	Parameter for denitrification rate in soil *denitr* (general*)	0.001-0.1 (1/d)
	Decay of humusN to fastN *degradhn* (land use dependent)	0.00001-0.1 (1/d)
	Mineralization of fastN to inorganicN *minerfn* (land use dependent)	0.000001-0.1 (1/d)
	Fraction of nutrient uptake in the uppermost soil layer *uptsoil1* (land use dependent)	0.001-1.0 (-)
	Number of days that fertilizer applications occur *fertdays* (general*)	10-150 (d)
Nitrogen processes in stream	Production/decay of N in water *wprodn* (general*)	0.0001-0.1 ($kg/m^3/d$)
	Parameter for denitrification in water *denitw* (general*)	0.000001-0.1 ($kg/m^2/d$)
	Parameters for calculation of water velocity in watercourses *rivvel1*, *rivvel2*, *rivvel3*	0.01-1.0 (-)

Table 2
Click here to download Table: Table 2.docx

Table 2. Calibrated nitrogen-export process parameters and their prior and posterior statistics resulting from calibration to either fortnightly or daily nitrate datasets. Statistics include prior uncertainty intervals, Relative Composite Sensitivity (RCS), and posterior statistics (maximum likelihood value (MAP), average, and standard deviation (Std)).

Parameter	Physical meaning	Prior	Posterior statistics							
			Fortnightly nitrate dataset				Daily nitrate dataset			
			RCS	MAP	Average	Std	RCS	MAP	Average	Std
denitr	Parameter for denitrification rate in soil (1/d)	0.001-0.1	0.0043	0.029	0.035	0.0088	0.0086	0.027	0.026	0.0013
denitw	Parameter for denitrification in water (kg/m^2/d)	10^{-6}-0.1	5.6×10^{-7}	9.1×10^{-4}	7.6×10^{-4}	5.2×10^{-4}	1.8×10^{-6}	9.5×10^{-4}	9.7×10^{-4}	7.8×10^{-5}
wprodn	Production/decay of N in water (kg/m^3/d)	10^{-4}-0.1	1.6×10^{-7}	0.0048	0.0057	0.0035	1.7×10^{-6}	0.0040	0.0041	6.3×10^{-4}
uptsoil102	Fraction of nutrient uptake in the uppermost soil layer for arable land (-)	0.001-1	0.012	0.999	0.97	0.042	0.026	0.999	0.997	0.0036
uptsoil107	Fraction of nutrient uptake in the uppermost soil layer for coniferous forest (-)	0.001-1	0.0029	0.39	0.53	0.27	0.0052	0.0023	0.020	0.029
uptsoil108	Fraction of nutrient uptake in the uppermost soil layer for mixed forest (-)	0.001-1	6.3×10^{-5}	0.95	0.64	0.22	2.8×10^{-4}	0.96	0.96	0.010
rivvel2	Parameters for calculation of water velocity in watercourses (-)	0.01-1	1.1×10^{-9}	0.22	0.48	0.26	2.6×10^{-5}	0.22	0.22	0.0040
fertdays	Number of days that fertilizer applications occur (d)	10-150	0.0052	85	93	20	0.0089	78	78	1.9

Table 3. Statistical model performance, in terms of streamflow at Silberhuette, Meisdorf and Hausneindorf gauging stations, during 1st Nov 2010 - 31st Oct 2013 and 1st Nov 2013 - 31st Oct 2015. *PBIAS* and *MAE* have the units of % and m^3/s, respectively, while *NSE* and *RSR* are unitless.

Sub-basin	1st Nov 2010 - 31st Oct 2013				1st Nov 2013 - 31st Oct 2015			
	NSE	*PBIAS*	*MAE*	*RSR*	*NSE*	*PBIAS*	*MAE*	*RSR*
Silberhuette	0.89	-2.9	0.26	0.33	0.70	6.5	0.23	0.55
Meisdorf	0.78	-2.5	0.49	0.47	0.74	-10.2	0.37	0.51
Hausneindorf	0.88	5.2	0.47	0.35	0.71	-11.6	0.42	0.54

Table 4
Click here to download Table: Table 4.docx

Table 4. Statistical model performance in terms of stream nitrate concentrations and nitrate loads for calibration (1st Nov 2010 - 31st Oct 2013) and validation (1st Nov 2013 - 31st Oct 2015) periods. Results are given for the three gauging stations (Silberhuette, Meisdorf and Hausneindorf), and as obtained from calibration using fortnightly and daily nitrate datasets. The columns “Fortnightly” and “Daily” refer to the results of models calibrated to, respectively, fortnightly and daily nitrate datasets. The goodness-of-fit statistics refer to model-measurement comparisons based on daily nitrate datasets (i.e., daily average stream nitrate concentrations, and daily nitrate loads estimated as the product of streamflow and nitrate concentration).

Variable	Criterion	Calibration						Validation					
		Silberhuette		Meisdorf		Hausneindorf		Silberhuette		Meisdorf		Hausneindorf	
		Fortnightly	Daily	Fortnightly	Daily	Fortnightly	Daily	Fortnightly	Daily	Fortnightly	Daily	Fortnightly	Daily
Daily average nitrate concentr-ations	*NSE*	0.57	0.68	0.43	0.52	-0.95	-0.56	0.74	0.76	0.61	0.66	-8.55	-5.53
	PBIAS (%)	17.3	4.75	11.2	-4.73	7.81	-1.25	4.60	-8.02	15.1	-3.09	58.4	43.9
	MAE (mg/L)	0.63	0.52	0.69	0.62	0.92	0.79	0.47	0.43	0.53	0.46	1.32	1.07
	RSR	0.65	0.57	0.75	0.70	1.40	1.25	0.51	0.49	0.63	0.58	3.09	2.55
Daily nitrate loads	*NSE*	0.77	0.75	0.54	0.50	0.71	0.78	0.67	0.74	0.76	0.69	0.29	0.46
	PBIAS (%)	-8.56	-19.6	-19.3	-32.2	22.7	13.6	3.61	-10.2	-11.5	-25.4	48.8	34.9
	MAE (kg/d)	92.4	92.4	193	204	262	225	69.4	58.5	87.6	91.5	203	172
	RSR	0.48	0.50	0.68	0.71	0.54	0.47	0.58	0.51	0.49	0.55	0.84	0.74

Table 5. Comparison of 95% prediction uncertainties in simulated stream nitrate concentrations at Silberhuette during 1st Nov 2010 - 31st Oct 2013, estimated from calibration using fortnightly and daily nitrate datasets. Par-Unc represents 95% parametric prediction uncertainty intervals of nitrate concentrations (mg/L); Tot-Unc represents 95% total prediction uncertainty intervals of nitrate concentrations (mg/L).

Criteria	Fortnightly nitrate dataset		Daily nitrate dataset	
	Par-Unc	Tot-Unc	Par-Unc	Tot-Unc
ARIL	0.97	9.2	0.15	4.7
PCI (%)	25	100	8.1	97
PUCI	0.31	0.10	0.87	0.21

Figure 1
Click here to download high resolution image

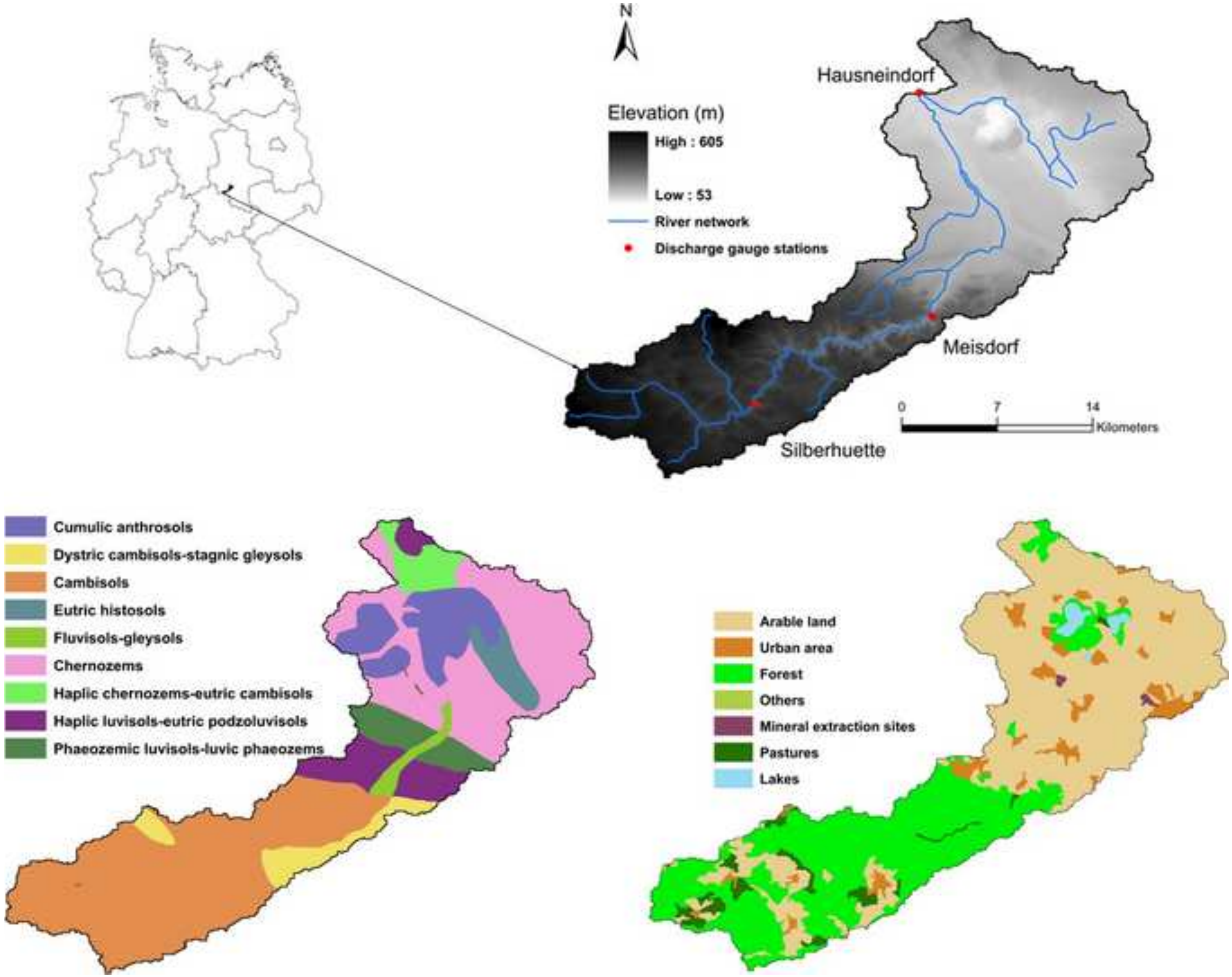

Figure 2

Click here to download high resolution image

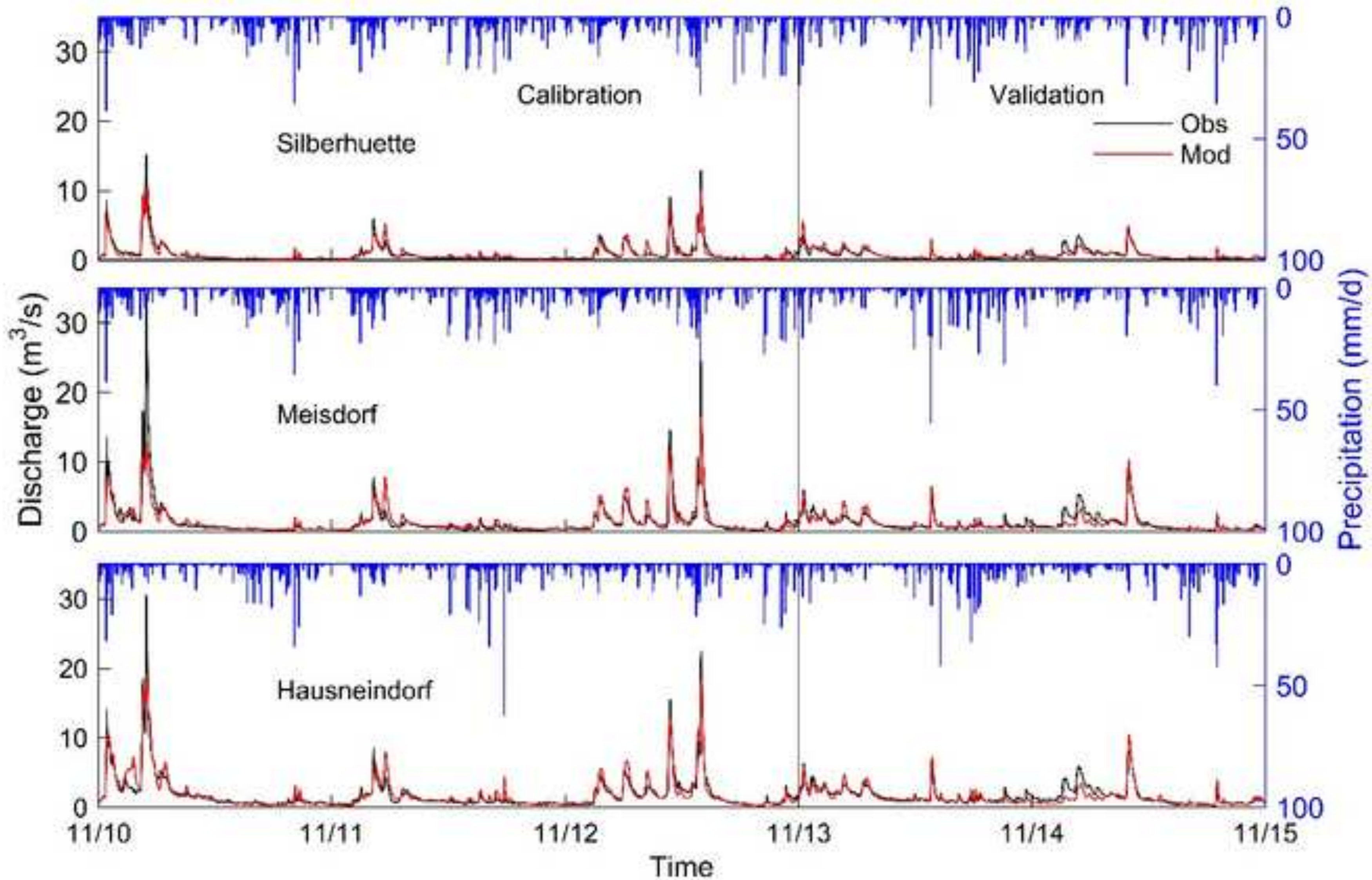

Figure 3
Click here to download high resolution image

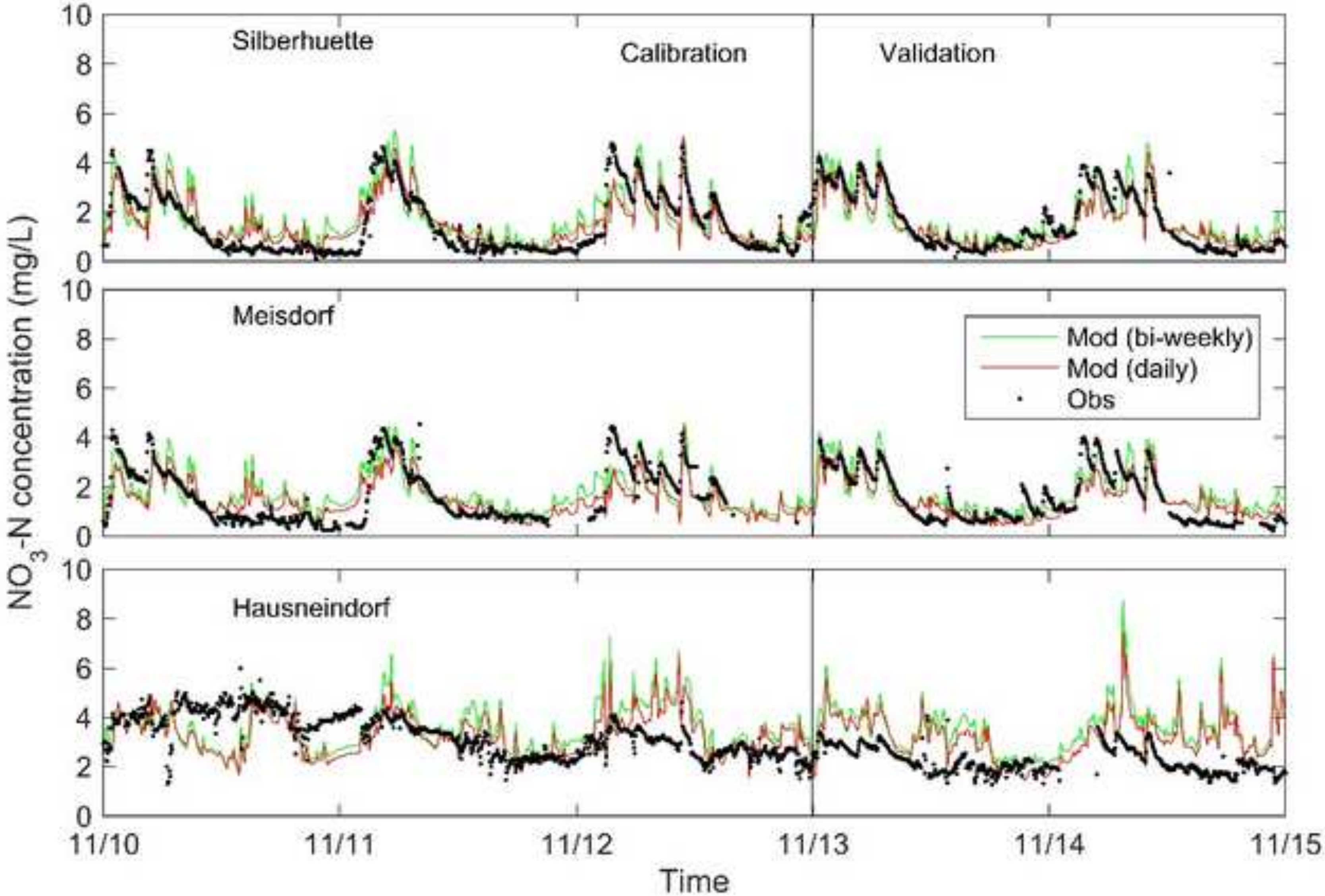

Figure 4
Click here to download high resolution image

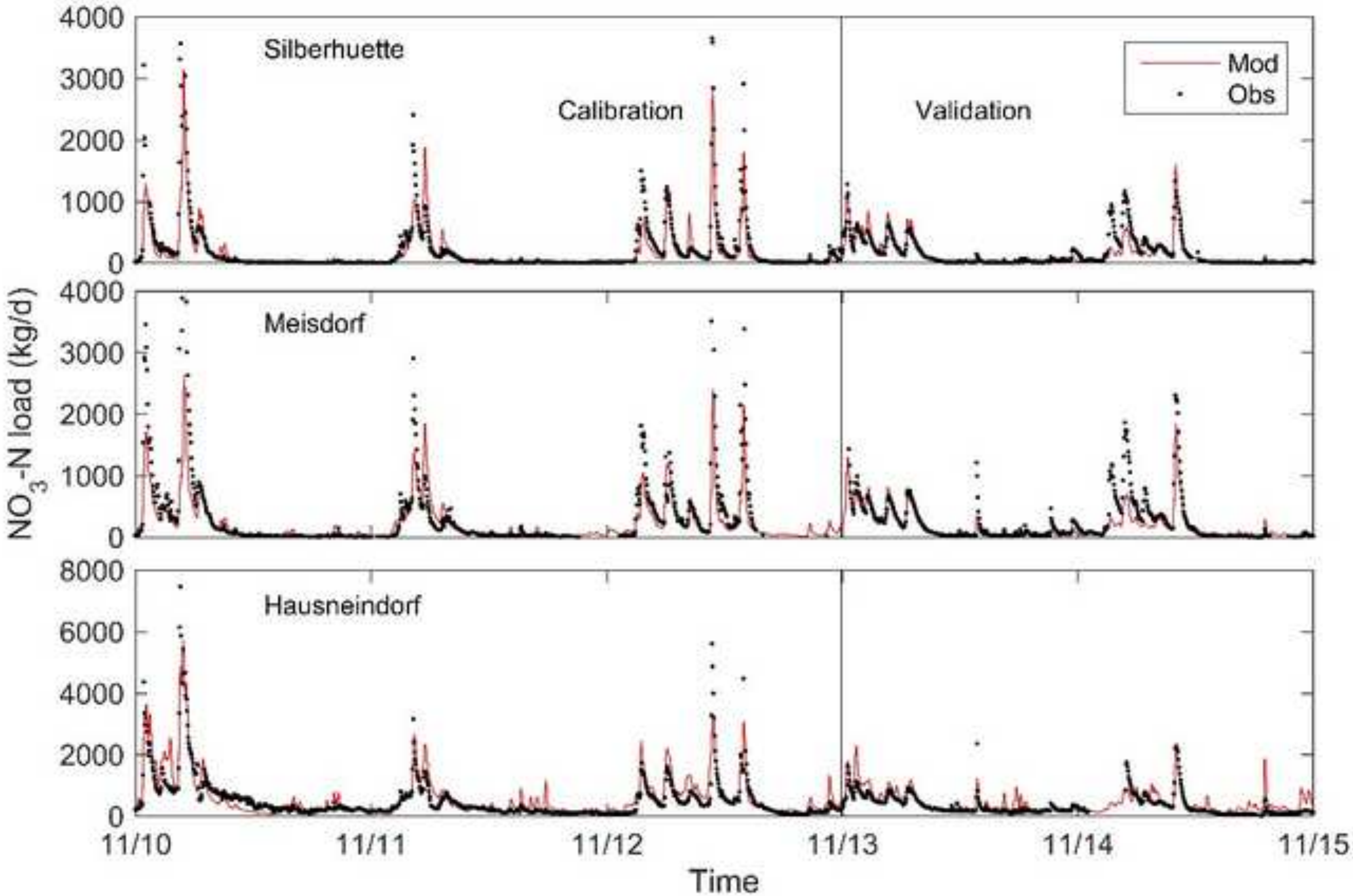

Figure 5

Click here to download high resolution image

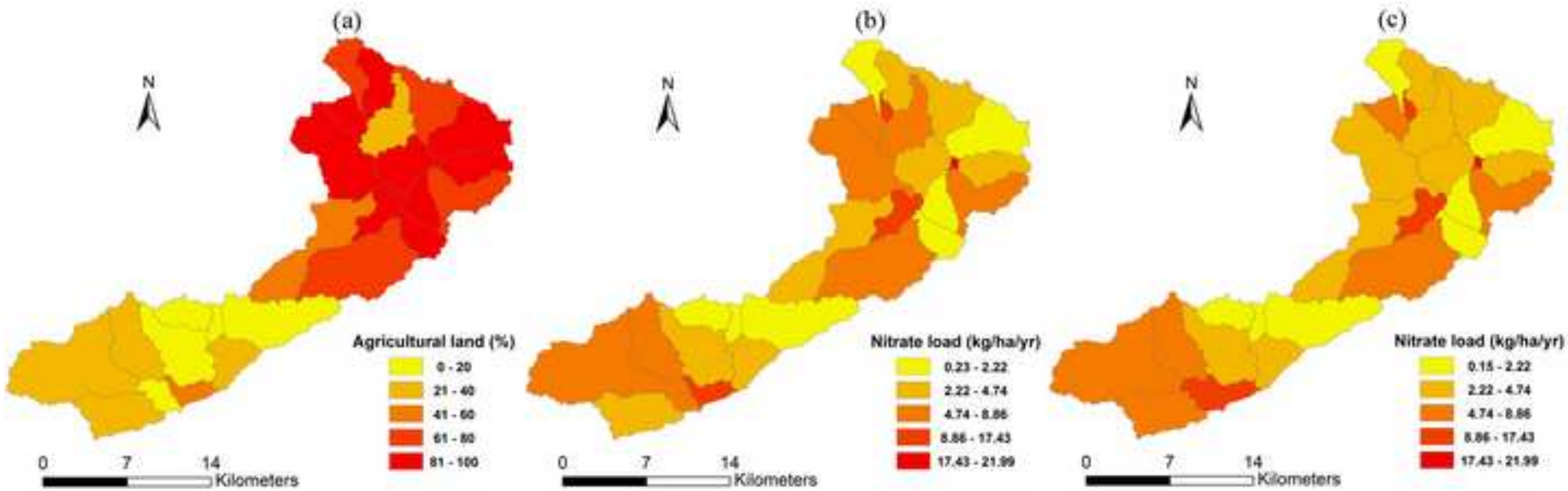

Figure 6
Click here to download high resolution image

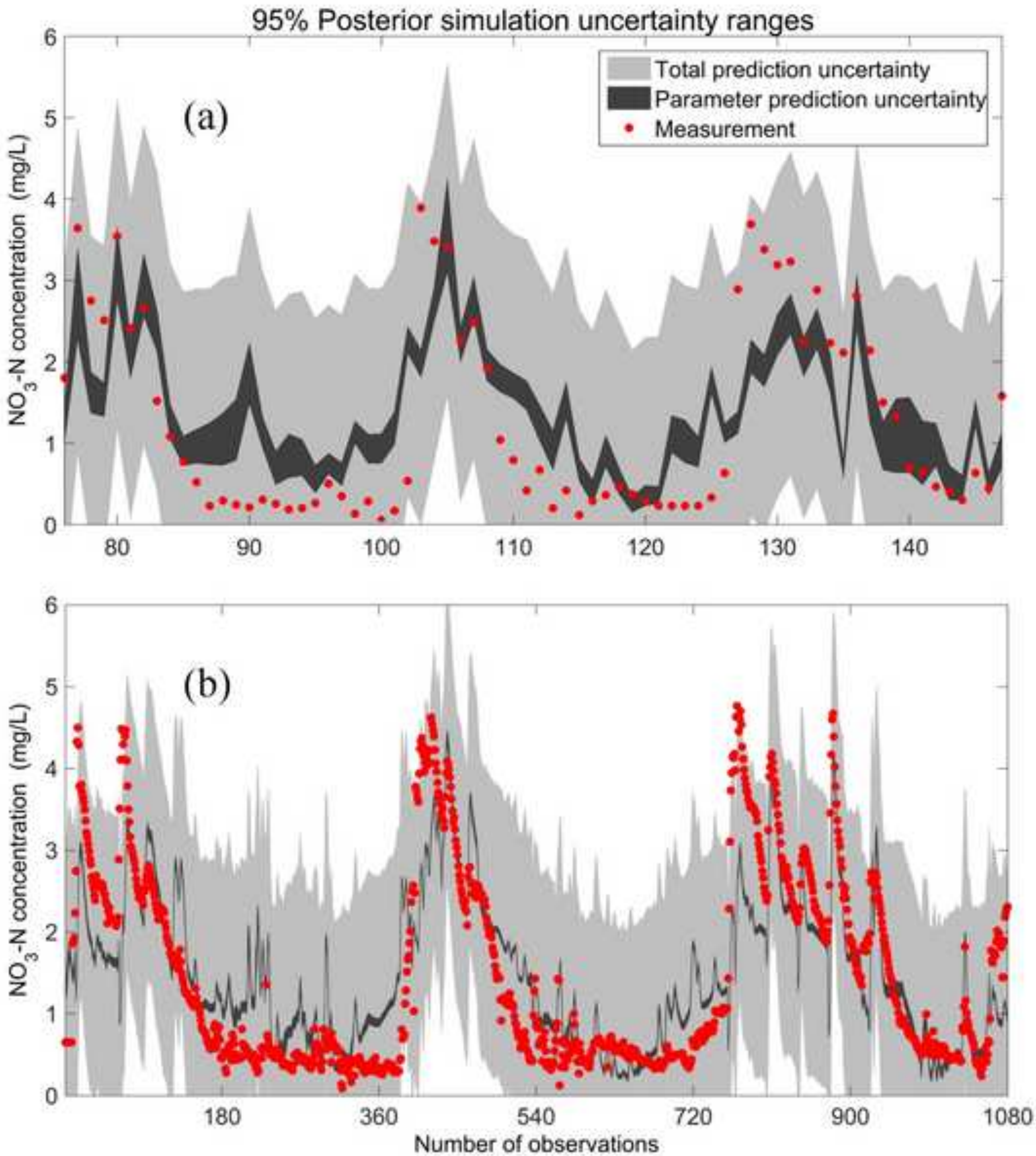

Figure captions

Figure 1. The Selke catchment: (a) Digital Elevation Model (DEM) and locations of streamflow and water quality gauging stations, (b) soil types, and (c) land use.

Figure 2. Streamflow simulation results together with observed daily rainfall during the periods of 1st Nov 2010 - 31st Oct 2013 (calibration) and 1st Nov 2013 - 31st Oct 2015 (validation) at Silberhuette, Meisdorf and Hausneindorf discharge gauging stations.

Figure 3. Observed (Obs) and Simulated (Mod) stream nitrate concentrations during calibration (1st Nov 2010 - 31st Oct 2013) and validation (1st Nov 2013 - 31st Oct 2015) periods. Results are shown for the three gauging stations (Silberhuette, Meisdorf and Hausneindorf), and as obtained from calibration using fortnightly and daily nitrate datasets.

Figure 4. Observed (Obs) and simulated (Mod) daily nitrate loads during calibration (1st Nov 2010 - 31st Oct 2013) and validation (1st Nov 2013 - 31st Oct 2015) periods at the three gauging stations of Silberhuette, Meisdorf and Hausneindorf. “Observed daily nitrate loads” are nitrate loads calculated as the product of observed streamflow and nitrate concentration. “Simulated daily nitrate loads” represent the nitrate loads estimated using simulated streamflow and nitrate concentration from the model calibrated to daily data.

Figure 5. Simulated time- and area-averaged nitrate loads (kg/ha/yr) at Selke catchment during 1st Nov 2010 - 31st Oct 2015 following calibration using daily and fortnightly nitrate datasets. (a) percentage of agricultural land; (b) average nitrate loads following calibration against daily nitrate data; (c) average nitrate loads following calibration against fortnightly nitrate data.

Figure 6. Comparison of 95% prediction uncertainty intervals of nitrate concentrations at Silberhuette during the period 1st Nov 2010 - 31st Oct 2013, estimated from calibration using: (a) fortnightly, and (b) daily nitrate datasets. Black bands represent parametric prediction uncertainty intervals, grey bands represent total prediction uncertainty intervals, and red dots represent the corresponding stream nitrate measurements.